Utrecht University Library
Partner in Science
Universiteit Utrecht

Pedagogische studiën 1990 : Tijdschrift voor onderwijskunde en opvoedkunde

https://hdl.handle.net/1874/205475

Sa oct. 6441

Pedagogische Studiën

Tijdschrift voor onderwijskunde en opvoedkunde

Inhoud Jaargang 67 1990

RIJKSUNIVERSITEIT TE UTRECHT
2207 6248

BIBLIOTHEEK DER RIJKSUNIVERSITEIT UTRECHT

Wolters-Noordhoff

Bijlage bij: *Pedagogische Studiën*, No. 1, januari 1991

Pedagogische Studiën

Tijdschrift voor onderwijskunde en opvoedkunde

REDACTIE

Prof. dr. J.Lowyck (voorzitter)
Dr. N. Verloop (secretaris)
Prof. dr. L.J.Th. van der Kamp
Prof. dr. L.F.W. de Klerk
Prof. dr. A.M.P. Knoers
Prof. dr. A. van der Leij
Prof. dr. C.F.M. van Lieshout
Prof. dr. G.W. Meijnen
Prof. dr. W.J. Nijhof
Dr. J.H. Slavenburg

CORRESPONDENTEN

Prof. dr. G. Kanselaar
Dr. A.M.J. Kloosterman
Prof. dr. E.C.H. Marx
Prof. dr. J.M. van Meel
Prof. dr. W. Wielemans

Redactiesecretariaat: Mevr. C. Hamburg, Postbus 251, 6800 AG Arnhem

Uitgever: Wolters-Noordhoff bv, Postbus 567, 9700 AN Groningen, tel. 050-22 68 86.
Correspondentie over abonnementen en losse nummers uitsluitend aan de uitgever.

Inhoudsopgave
Jaargang 67 1990

AUTEURSREGISTER

ARTIKELEN:

Baltussen, M. W. J. & E. C. D. M. van Lieshout
Rekenfeiten in het cijferend optellen en aftrekken tot 100 306

Bergh, H. van den & H. Kuhlemeier
De haalbaarheid van eindtermen Nederlands voor de basisvorming 1

Bon, W. H. J. van, M. H. G. Dierx & J. J. M. P. Klerkx
De morfologische competentie van kinderen met leerproblemen 59

Bon, W. H. J. van & F. G. M. Bouwens
Brugklassers met discordante lees- en spelprestaties 317

Bos, K. Tj., A. M. Ruijters & A. J. Visscher
Absenteïsme in het voortgezet onderwijs 341

Bouwens, F. G. M.
zie: Bon, W. H. J. van

Corte, E. De, L. Verschaffel & H. Schrooten
Cognitieve effecten van het systematisch leren programmeren in LOGO een onderwijsexperiment bij zesdeklassers 293

Daal, V. H. P. van
zie: Leij, A. van der

Damme, J. Van
zie: Vanderlocht, M.

Dierx, M. H. G.
zie: Bon, W. H. J. van

Driessen, G
zie: Jungbluth, P.

Gerris, J. R. M.
zie: Janssens, J. M. A. M.

Hulst, J. W. van
Martinus J. Langeveld (1905-1989) ... 158

Janssen, A. W. H.
zie: Janssens, J. M. A. M.

Janssens, J. M. A. M., J. R. M. Gerris & A. W. H. Janssen
Ouderlijke opvoedingsgedragingen, empathie en prosociale ontwikkeling van het kind 403

Jongeling, E.
zie: Richardson, R.

Jungbluth, P., Chr. Verhaak & G. Driessen
Vervolgadviezen in relatie tot etniciteit; hoe verhouden prestaties, gezinsachtergrond en leerkrachtoordelen zich tot elkaar? 231

Klerkx, J. J. M. P.
zie: Bon, W. H. J. van

Kloeze, J. W. te
zie: Wouw, B. A. J. van der

Kuhlemeier, H
zie: Bergh, H. van den

Kuyk, J. J. van
Kunnen jonge kinderen 'Ordenen'? Onderzoek naar deelvaardigheden voorbereidend rekenen van 4 – 6 jarigen 429

Leenders, F. J.
zie: Stokking, K. M.

Leij, A. van der, H. Smeets & V. H. P. van Daal
De ontwikkeling van de Computergestuurde Toets voor Automatisering van het Lezen 163

Lem, P
zie: Veenman, S.

Lieshout, E. C. D. M. van
zie: Baltussen, M. W. J.

Lil, J. E. van
zie: Nikken, P

Lowyck, J.
zie: Vanmaele, L.

Meijers, F. & R. van Wijk
'Geef mij maar werk'. Jongeren over de tijdelijke voorziening gemeentelijke werkgelegenheidsinitiatieven voor jongeren 16

Nijhof, W. J.
Vliegwieleffecten van een stimuleringsprogramma. Een evaluatie van INSP-deelprogramma's gericht op het beroepsgeoriënteerde onderwijs 100

In memoriam Professor Dr. M. J. Langeveld 157

Nijhof, W. J. & J. L. M. Remmers
Basisvaardigheden en het beroepsonderwijs. Een operationalisering van het begrip basisvaardigheden 389

Nijhof, W. J. & J. Scheerens
Ten geleide Themanummer Evolutie in onderwijsevaluatie 85

Nijhof, W. J. & J. Scheerens
De progressie in de Nederlandse evaluatiepraktijk 151

Nikken, P. & J. E. van Lil
Determinanten van TV-opvoeding: een replicatie-onderzoek 417

Peschar, J. L.
Mèt of zonder basisvorming? Evaluatie in het voortgezet onderwijs 116

Reezigt, G. J.
zie: Weide, M. G.

Remmers, J. L. M.
zie: Nijhof, W. J.

Richardson, R., E. Jongeling & F. Wester
Systematisering van de analyse in de gefundeerde theoriebenadering; een illustratie 355

Roelofs, E.
zie: Veenman, S.

Ruijters, A. M.
zie: Bos, K. Tj.

Schaper, M. W.
Het effect van articulatie op het onthouden van geschreven woorden bij dove kinderen 26

Scheerens, J.
De evaluatie van het onderwijsdeel van het Informatica Stimuleringsplan (INSP) 88

Scheerens, J.
zie: Nijhof, W. J.

Schrooten, H.
zie: Corte, E. De

Smeets, H
zie: Leij, A. van der

Stokking, K. M. & F. J. Leenders
Evaluatie van een beleidsmaatregel. Uitvoering en effecten van de nascholing zorgverbreding lezen 245

Terlouw, C.
Leren denken binnen het vakgebied bestuurskunde: het leren oplossen van bestuurskundige ontwerpproblemen . 179

Tesser, P.
zie: Werf, M. P. C. van der

Vanderlocht, M. & J. Van Damme
Kennis en het oplossen van natuurkundige problemen. Enkele aandachtspunten voor de onderwijspraktijk 70

Vanmaele, L. & J. Lowyck
Het schrijfmodel van Hayes en Flower: een kritische analyse 205

Veenman, S., P. Lem & E. Roelofs
Omgaan met combinatieklassen: een onderzoek naar de effecten van een teamgericht nascholingsprogramma .. 45

Verhaak, Chr.
zie: Jungbluth, P.

Verhetsel, A.
De wereld in ons hoofd. Een onderwijsexperiment in verband met het ruimtelijke voorstellings- en structureringsvermogen 261

Verschaffel, L.
zie: Corte, E. De

Visscher, A. J.
zie: Bos, K. Tj.

Weide, M. G. & G. J. Reezigt
Differentiatie en selectie in het basisonderwijs 222

Werf, M. P. C. van der & P. Tesser
De evaluatie van het onderwijsvoorrangsbeleid 127

Wester, F.
zie: Richardson, R.

Wijk, R. van
zie: Meijers, F.

Wijnstra, J. M.
Periodiek peilingsonderzoek: de opzet van het project PPON geïllustreerd aan de rekenpeiling einde basisonderwijs 139

Wouw, B. A. J. van der & J. W. te Kloeze
Kinderen en autonomie in de vrije tijd 368

DISCUSSIE-BIJDRAGEN:

Dronkers, J.
Kanttekeningen bij het congres ter afsluiting van het Rotterdamse project Onderwijs en Sociaal Milieu (1969-1988) 193

Harskamp, E. & C. Suhre
Kritiek op methodenonderzoek tussen wal en schip 284

Heuvel-Panhuizen, M. van den
Rekenmethoden-onderzoek op de proef gesteld 271

Scheerens, J.
Het koekoeksjong uit het ei van Columbus. Commentaar op het rapport "Schaal en kwaliteit in het basisonderwijs". 444

Slavenburg, J. H.
Criteria en evaluatie, een moeilijke combinatie 196

Suhre, C.
zie: Harskamp, E.

KRONIEKEN:

Leij, A. van der
PEDON-Conferentie 'Het oog op agogisch onderzoek' 326

Nijhof, W. J.(red.)
De AERA Annual Meeting: Boston 1980, met bijdragen van *E. van den Berg, G. J. J. Biesta, K. Bügel, W. van de Grift, J. F. M. J. van Hout, L. F. W. de Klerk, S. Miedema, M. Mulder, W. J. Nijhof, H. Robben, P. Sanders, D. Thio, Th. Wubbels* 450

Veer, R. van der
Onderwijsontwikkelingen in de USSR 82

BOEKBESPREKINGEN:

Asselbergs-Neessen, V.
Kind, kunst en opvoeding. De Nederlandse beweging voor beeldende expressie (*A. Th. van Meel-Jansen*) 384

Baudoin, T., T. Meelis-Voorma, P. Molenaar, H. Overmeijer & B. Ros (red.)
Jeugdliteratuur voor de beroepspraktijk. Een leergang voor het HBO (*A. W. M. Duijx*) 475

Bergh, H. van den
Examens geëxamineerd. Een onderzoek naar de validiteit van examens Nederlands voor LBO en MAVO (*L. J. Th. van der Kamp*) 330

Brouwer, C. N.
Geïntegreerde lerarenopleiding; principes en effecten (*J. A. de Jong*) 336

Brunschot, R. A. M. van
Begeleiding van beginnende docenten in het technisch onderwijs (*W. Jochems*) 471

Goffree, F. & K. Buys
Tegengesteld, wiskundedidactiek op de grens van basis- en voortgezet onderwijs, toegespitst op negatieve getallen (*J. M. C. Nelissen*) 40

Graaf, W. A. W. de
De zaaitijd bij uitnemendheid; jeugd en puberteit in Nederland (*J.C. Sturm*) 338

Graaff, E. de
Simulation of Initial Medical Problemsolving (*H. Verstralen*) 385

Karstanje, P. N.
Beleidsevaluatie bij controversiële onderwijsvernieuwing (*J. Scheerens*) 199

Koops, W. & J. J. van der Werff (red.)
Overzicht van de empirische ontwikkelingspsychologie. 1. Grondkenmerken van het vak (*E. Elbers*) 155

Koops, W. & J. J. van der Werff(red.)
Overzicht van de empirische ontwikkelingspsychologie. 3. Sociale ontwikkeling en de ontwikkeling van de persoonlijkheid (*J. L. van der Linden*) ... 382

Lumer-Henneböle, B. & E. Nijssen
Basisschulen in den Niederlanden (*W. Wielemans*) 201

Mulder, E.
Beginsel en Beroep. Pedagogiek aan de Universiteit in Nederland 1900-1940 (*B. Levering*) 333

Sanden, J. M. M. van der (red.)
Huiswerk en huiswerkbeleid (*A. M. P. Knoers*) 154

Singer, E.
Kinderopvang en de moeder-kindrelatie (*M. Lunenberg*) 427

Singley, M. K. & J. R. Anderson
The transfer of cognitive skill (*J. H. Slavenburg*) 475

Slavenburg, J. H. & T. A. Peters (red.)
Het project Onderwijs en Sociaal Milieu: een eindbalans (*P. G. Swanborn*) 470

Sturm, J. C.
Een goede gereformeerde opvoeding (*J. D. Imelman*) 473

Swaan, A. de
Zorg en de staat. Welzijn, onderwijs en gezondheidszorg in Europa en de Verenigde Staten in de nieuwe tijd (*M. Santema*) 201

Teyken, C.
Ontwerpen als proces (*C. Terlouw*) ... 238

Tomic, W. & P. C. van der Sijde
Changing teaching for better learning (*H. H. Tillema*) 472

Verloop, N.
Interactive cognitions of student-teachers. An intervention study (*A. H. Corporaal*) 332

Voogt, J. C.
Scholen doorgelicht. Een studie over schooldiagnose (*J. Scheerens*) 383

Vos, J.
Meer dan feiten... Werken met het informatieve kinderboek op de basisschool (*A. W. M. Duijx*) 39

Zuidema, J.
Efficiënt spellingonderwijs. Een leer- en expertmodel voor het spellen (*M. J. C. Mommers*) 38

THEMANUMMER:

Evolutie in onderwijsevaluatie (maartnummer)

W. J. Nijhof & J. Scheerens
Ten geleide 85

J. Scheerens
De evaluatie van het onderwijsdeel van het Informatica Stimuleringsplan (INSP) 88

W. J. Nijhof
Vliegwieleffecten van een stimuleringsprogramma. Een evaluatie van INSP-deelprogramma's gericht op het beroepsgeoriënteerde onderwijs 100

J. L. Peschar
Mèt of zonder basisvorming? Evaluatie in het voortgezet onderwijs 116

M. P. C. van der Werf & P. Tesser
De evaluatie van het onderwijsvoorrangsbeleid 127

J. M. Wijnstra
Periodiek peilingsonderzoek: de opzet van het project PPON geïllustreerd aan de rekenpeiling einde basisonderwijs 139

W. J. Nijhof & J. Scheerens
De progressie in de Nederlandse evaluatiepraktijk 151

652 TS oct 6441

De haalbaarheid van eindtermen Nederlands voor de basisvorming

H. VAN DEN BERGH
Rijksuniversiteit Utrecht
H. KUHLEMEIER
Cito, Arnhem

Samenvatting

Eindtermen in de basisvorming worden geacht twee functies te vervullen. Enerzijds fungeren eindtermen als drempelniveau: ze beschrijven het 'minimumniveau' dat in principe door iedere leerling bereikt moet worden. Anderzijds vormen ze een richtpunt: ze beschrijven het 'streefniveau' waar alle docenten en leerlingen naar toe moeten werken. In dit artikel wordt het vraagstuk van de verenigbaarheid van beide functies aangesneden. Op basis van empirische gegevens uit een recent uitgevoerd peilingsonderzoek wordt aannemelijk gemaakt dat men over de hoogte van het door vrijwel iedere vijftien- à zestienjarige behaalde onderwijsniveau geen al te hoge verwachtingen mag koesteren. Daarmee rijst de vraag naar de mogelijkheid en de zinvolheid van het nastreven van 'één minimumniveau voor allen'. Dit vraagstuk wordt ook bezien in het licht van de verschillen tussen leerlingen in leerprestaties.

Bij het formuleren van eindtermen en de constructie van toetsen ter afsluiting van de basisvorming doen zich twee, nog niet opgeloste problemen voor. Het eerste heeft betrekking op de normstellingsproblematiek. Hoe ontwikkelen we algemeen geaccepteerde en realistische minimumeisen? Het tweede probleem betreft de toepassing van de normen. Hoe beslissen we of een leerling het niveau van de basisvorming bereikt heeft? In dit artikel worden drie methoden om normen te ontwikkelen gepresenteerd en aan de hand van concrete leerlingprestaties uit peilingsonderzoek toegepast.

1 *Inleiding*

De invoering van de basisvorming is ontegenzeglijk de grootste operatie in het voortgezet onderwijs sinds de Mammoetwet. Door de val van het kabinet Lubbers II is het doorgaan van de plannen weliswaar nog onzeker geworden, maar dit neemt niet weg dat de discussie over de zin of onzin van de operatie nog steeds onverminderd voortwoedt. De verwachtingen omtrent het effect van de basisvorming zijn hoog gespannen. Met de invoering van de basisvorming beoogt de overheid een algehele verhoging van het onderwijsniveau te bewerkstelligen, de verschillen tussen leerlingen te verkleinen, het keuzemoment uit te stellen, de inhoud van het onderwijs beter controleerbaar te maken voor parlement en samenleving, de aansluiting op het vervolgonderwijs te verbeteren et cetera (Ministerie van Onderwijs en Wetenschappen, 1989). Daarnaast wordt een omvangrijke vernieuwing op het terrein van de vakinhouden, de vakdidactiek en de toetsing nagestreefd (Rijlaarsdam, 1989a; 1989b; Ten Brinke, 1989; Wijgh, 1989a). Zo adviseerde de Eindtermencommissie Nederlands (1989) voor de vakonderdelen lezen, schrijven, spreken, luisteren en taalbeschouwing een groot aantal eindtermen. Daarin is vastgelegd wat de leerlingen aan het einde van de basisvorming ten minste moeten kennen en kunnen. Daarnaast worden voor het vak Nederlands diverse verschuivingen in het onderwijsaanbod geadviseerd. Bij voorbeeld: meer aandacht voor de mondelinge taalvaardigheid en een grotere gerichtheid op het gebruik van taal in communicatieve situaties. Naar het zich laat aanzien zal de invoering van de basisvorming voor de scholen in het voortgezet onderwijs drastische veranderingen met zich meebrengen. De meeste scholen zullen hun onderwijs op twee niveaus moeten aanbieden: een algemeen niveau dat voor tachtig procent van de leerlingen haalbaar moet zijn en een hoger niveau waaraan 40 procent van hen moet voldoen. Ook voor de leerlingen zijn de consequenties groot. Om tot het vierde leerjaar VWO toegelaten te worden

dienen negen vakken op het hogere niveau te worden afgesloten; voor de doorstroming naar HAVO/MBO geldt de eis van zeven vakken op hoger niveau.

In het kader van de basisvorming speelt de problematiek van het ontwikkelen en toepassen van normen een essentiële rol. Immers, een wezenlijk element en misschien wel de kern van de basisvorming, vormen de eindtermen. Wat moet een leerling minimaal kennen en kunnen? Het formuleren van eindtermen is geen sinecure. Ze moeten aan diverse eisen voldoen: duidelijk, algemeen acceptabel, onderwijsbaar, toetsbaar et cetera. Bovendien moeten ze voor de leerlingen zowel haalbaar als uitdagend zijn. Dit betekent dat ze noch te hoge noch te lage eisen aan de leerlingen mogen stellen. Eindtermen mogen met name niet al te ver vooruitlopen op de gangbare onderwijspraktijk en het feitelijke prestatieniveau van de leerlingen. Moeilijk onderwijsbare eindtermen zullen de onrust onder docenten alleen nog maar verhogen. Van te hoog gegrepen eindtermen zal voor veel leerlingen een demotiverende werking uitgaan ('dat is voor mij toch te moeilijk'). Van de andere kant zullen te lage eisen de motivatie eveneens niet bevorderen, aangezien ze zowel voor docenten als leerlingen geen uitdaging vormen ('dat is zo gemakkelijk, daar hoef ik me toch niet voor in te spannen').

Met betrekking tot de niveauproblematiek bezitten eindtermen twee functies. Enerzijds fungeren eindtermen als drempelniveau, dit wil zeggen: ze beschrijven het *minimum*niveau dat in principe door iedere leerling gehaald moet worden. Anderzijds dienen eindtermen als richtpunt: ze beschrijven het *streef*niveau voor alle docenten en leerlingen (Wijgh, 1989b). Men kan zich afvragen in hoeverre deze twee functies met elkaar verenigbaar zijn. Hoe zinvol zijn minimumeisen als streefniveau? Deze vraag naar de zinvolheid van het nastreven van 'één niveau voor allen' kan natuurlijk pas beantwoord worden wanneer we weten hoe hoog het prestatieniveau is dat feitelijk door iedere vijftien- à zestienjarige gehaald wordt. Is dit niveau laag, dan zal er van de eindtermen die dit niveau specificeren voor veel docenten en leerlingen nauwelijks een inspirerende werking kunnen uitgaan. Dit geldt in nog sterkere mate wanneer we bedenken dat leerlingen qua prestatieniveau vaak sterk verschillen. Over de hoogte van dit communale niveau zijn de meningen verdeeld. Met betrekking tot de leerprestaties voor het vak Nederlands zijn sommigen de mening toegedaan dat hetgeen door werkelijk alle leerlingen beheerst wordt niet veel meer omvat dan het kunnen opzeggen van het alfabet. Volgens anderen zou het communale minimumniveau aanzienlijk hoger (kunnen) liggen. Zo formuleerde de Eindtermencommissie Nederlands (1989) voor de vakonderdelen lezen, schrijven, spreken, taalbeschouwing en informatievaardigheden 30 concrete eindtermen. Eén van de problemen bij het ontwikkelen van realistische doelstellingen is dat over het onderwijsaanbod en het feitelijke prestatieniveau van leerlingen maar weinig bekend is. Een inschatting van het aanbod en de leerprestaties onder de condities van de basisvorming is zo mogelijk nog moeilijker te geven. Dientengevolge is over de vraag naar de 'realiteitswaarde' van de thans voorliggende eindtermen maar weinig met zekerheid te zeggen. Indicatief is in dit verband dat geen van de 14 eindtermencommissies zich gewaagd heeft aan het geven van een voorspelling van het percentage leerlingen dat de basisvorming binnen de gestelde tijd zal bereiken. Het is dan ook niet verwonderlijk dat de haalbaarheid ervan door velen in twijfel wordt getrokken (ISCB, 1989; ARVO, 1989; PvdA, 1989).

Met dit artikel willen wij een empirische fundering geven aan de discussie over de haalbaarheid van eindtermen. Hiertoe worden de consequenties van verschillende normen gerelateerd aan de prestaties van derdeklassers uit de Proefpeiling Nederlands in het voortgezet onderwijs (Kuhlemeier & Van den Bergh, 1989b; Van den Bergh & Kuhlemeier, 1989). In dit artikel worden drie methoden om normen te ontwikkelen beschreven. Tevens wordt de toepassing ervan geïllustreerd aan de hand van gegevens uit de Proefpeiling Nederlands. Het artikel kent de volgende opbouw. Eerst worden de voor onze exploratie van de normstellingsproblematiek gebruikte PPON-taken beschreven en wordt een globale indruk gegeven van de prestaties van de leerlingen op deze taken (paragraaf 2). Vervolgens worden de drie door ons toegepaste methoden om normen te ontwikkelen beschreven en aan de hand van enkele PPON-taken toegepast (paragraaf 3). Daarbij wordt ook aandacht besteed aan een evaluatieve vergelijking van

de drie methoden. Het artikel besluit met enkele conclusies over de haalbaarheid van eindtermen Nederlands in de basisvorming (paragraaf 4).

2 *Prestaties voor enkele lees-, spreek- en schrijftaken*

In de Proefpeiling Nederlands in het voortgezet onderwijs is een veelheid aan opdrachten afgenomen bij een representatieve steekproef derdeklassers. De leesvaardigheid is gepeild met behulp van drie taakboekjes die ieder aan een aparte steekproef van circa 1450 derdeklassers zijn voorgelegd. Dit maakt het mogelijk om voor het vakonderdeel lezen drie relatief onafhankelijke schattingen van de normen te verkrijgen. De wijze waarop wij normen voor lezen ontwikkeld hebben wordt in dit artikel alleen geïllustreerd aan de hand van één van deze taken. Dit taakboekje, 'Schoolreis naar Waddenoog', bevat vijf[1] leesopdrachten. Daarbij moet de leerling zich verplaatsen in de situatie van een leerling die voor zijn klas een schoolreis naar het eiland Waddenoog organiseert. De eerste drie opdrachten kunnen als functioneel worden aangemerkt. Dit wil zeggen dat ze gebaseerd zijn op taalsituaties die ook buiten de school van belang kunnen zijn. Zo moet de leerling uit een door de V.V.V. verstrekte informatiefolder een kampeerboerderij zoeken die aan diverse eisen voldoet. Gevraagd wordt onder meer naar de openingstijden, het aantal slaapplaatsen, de aanwezigheid van een warmwatervoorziening, de ligging van de boerderijen en naar die éne boerderij die aan alle wensen voldoet. De twee andere opdrachten uit Schoolreis naar Waddenoog zijn 'traditionele' tekstbegriptoetsen, die qua thematiek nauw aansluiten op de functionele opdrachten in dit taakboekje.

De spreekvaardigheid is geëvalueerd met behulp van drie opdrachten die zijn voorgelegd aan ruim 1100 derdeklassers. De spreekopdracht 'Het navertellen van een verhaal' is een voorbeeld van een meer schoolse opdracht. Voor deze opdracht moet de leerling aan een medeleerling uitleggen hoe bakstenen gemaakt worden. Nadat gecontroleerd is of de leerling de uitleg op de band goed begrepen heeft, moet het proces aan een medeleerling uitgelegd worden. Hierbij kan de leerling gebruik maken van acht plaatjes die de verschillende stadia in het produktieproces weerspiegelen. De leerlingprestaties zijn beoordeeld op de aspecten inhoud alsmede taalgebruik en opbouw. De beide andere spreekopdrachten hebben een meer functioneel karakter (Blok, 1987). De leerling moet twee telefoongesprekken voeren met een bemiddelaar van Randstad Uitzendbureau (gespeeld door een telefoonassistent op het Cito). In het eerste telefoongesprek wordt de leerling vakantiewerk aangeboden (Randstad biedt aan). Getracht is het telefoongesprek zoveel mogelijk een natuurlijk karakter te geven. Vandaar dat regelmatig reacties aan de leerlingen ontlokt worden ('oké'; 'snap je?' e.d.). Op opmerkingen of vragen van de leerling moest vanzelfsprekend ingegaan worden. Na het telefoongesprek kreeg de leerling een situatieschets waaruit bleek dat aan vakantiewerk bij een uitzendbureau enkele vervelende consequenties verbonden kunnen zijn (Arbeidsvoorwaarden bij Randstad). De leerling moest wederom contact opnemen met de bemiddelaar van Randstad en de problemen rond vakantiewerk bevragen. Tijdens de telefoongesprekken mochten de leerlingen aantekeningen maken en vragen stellen over de in de telefoongesprekken verschafte informatie.

De wijze waarop de leerlingen de voorheen beschreven lees- en spreekopdrachten hebben

Tabel 1 *De gemiddelde prestatie als functie van de maximale score voor de opdrachten van 'Schoolreis naar Waddenoog' en de drie spreektaken*

Opdracht	Gemiddelde prestatie
Functionele leesopdrachten:	
1 lezen tabel en plattegrond	70%
2 invullen inschrijfformulier	75%
3 raadplegen bus- en bootdienstregeling	81%
Tekstbegriptoetsen:	
4 begrijpen biologietekst over duinen	68%
5 zoekend lezen in tekst over musea	81%
Spreektaken:	
1 Randstad biedt aan	85%
2 arbeidsvoorwaarden bij Randstad	60%
3 navertellen verhaal:	
inhoud	50%
taalgebruik en opbouw	63%

uitgevoerd is samengevat in Tabel 1. De moeilijkheid van de opdrachten is daarbij uitgedrukt als het percentage van de maximale score die de leerling voor de opdracht kon behalen. Tabel 1 laat zien dat de derdeklassers met de lees- en spreekopdrachten over het algemeen weinig moeite lijken te hebben. Wat betreft vijf van de zes leesopdrachten wekt dit weinig verwondering, omdat ze afkomstig zijn uit een eerder in het basisonderwijs uitgevoerde proefpeiling (Wesdorp, Van den Bergh, Bos, Hoeksma, Oostdam, Scheerens & Triesscheijn, 1985).

Het schrijfniveau van de derdeklassers is gepeild met behulp van vier schrijfopdrachten die ieder aan ruim 1450 derdeklassers zijn voorgelegd. De eerste schrijfopdracht is een functionele opdracht: het schrijven van een sollicitatiebrief. Op grond van een advertentie, waarin om een vakantiehulp gevraagd wordt, moet de leerling een sollicitatiebrief aan een groot warenhuis schrijven. Bij de tweede functionele opdracht moet de leerling een kort, formeel briefje schrijven aan de directeur van de school. De beide andere opdrachten zijn meer traditionele stelopdrachten. Bij de derde opdracht moet de leerling aan de hand van documentatiemateriaal een betoog schrijven voor de schoolkrant over de zin of onzin van huiswerk. De vierde en laatste opdracht omvat het schrijven van een persoonlijk getint verhaal voor de schoolkrant over een angstige ervaring.

De door de leerlingen geproduceerde teksten zijn beoordeeld op verschillende aspecten, zoals globale kwaliteit, inhoud, organisatie en stijl. Naast deze oordelen over vier aspecten van de schrijfprodukten is ook informatie verzameld over de mate waarin de leerling spellings- en grammaticale conventies correct gehanteerd heeft. Geteld zijn onder meer de aantallen spelfouten, fouten in de werkwoordspelling, interpunctiefouten, grammaticale en idiomatische fouten en het aantal welgevormde zinnen. Een indicatie van de prestaties op de schrijftaken is weergegeven in Tabel 2. Deze tabel behoeft enige toelichting. Omwille van een inzichtelijke rapportage is de score voor de algemene beoordelingsaspecten uitgedrukt op een schaal met een bereik van 1 (zeer slecht) tot 7 (zeer goed). Het aantal formele fouten dat een leerling maakt is afhankelijk van de lengte van de geschreven tekst. Daarom zijn deze aantallen fouten in Tabel 2 gestandaardiseerd.

Met betrekking tot de algemene aspecten van het schrijven kunnen we uit Tabel 2 concluderen dat de beide functionele opdrachten iets moeilijker lijken dan de twee conventionele stelopdrachten. Vooral de stijl van de sollicitatiebriefjes laat volgens de beoordelaars te wensen over; de inhoud van de excuusbriefjes is daarentegen wel goed verzorgd. Bij de technische aspecten valt op dat de leerlingen gemiddeld betrekkelijk weinig spelfouten maken en tamelijk weinig zondigen tegen de regels van de werkwoordspelling. Hierbij moet opgemerkt worden dat de gepresenteerde gemiddelde aantallen (werkwoord)spelfouten een enigszins vertekend beeld geven; de ver-

Tabel 2 *Prestaties op de schrijftaken voor enkele algemene en formele beoordelingsaspecten*

	Sollicitatie-brief	Excuus-brief	Betoog huiswerk	Verhaal angst
Algemene aspecten:				
globale kwaliteit	3.5	4.2	–	–
inhoud	4.5	6.3	–	4.1
publiekgerichtheid	–	–	3.9	–
doelgerichtheid	–	–	4.2	–
stijl	2.7	3.7	4.5	4.1
organisatie	4.0	4.3	4.4	4.1
Technische aspecten:				
aantal spelfouten per 100 woorden	4.6	2.7	1.8	1.6
aantal fouten in werkwoordspelling per 100 zinnen	2.4	3.5	6.1	4.4
aantal interpunctiefouten per 100 zinnen	96.9	80.9	71.9	59.1
aantal grammaticale en idiomatische fouten per 100 zinnen	44.1	41.9	43.5	29.9
aantal welgevormde zinnen per 100 zinnen	60.9	64.4	64.6	74.1

deling van de aantallen fouten is nogal scheef. Een klein deel van de leerlingen is namelijk 'verantwoordelijk' voor een groot deel van het absolute aantal fouten, terwijl de grote meerderheid weinig tot zeer weinig fouten maakt (Kuhlemeier & Van den Bergh, 1989b). Wel maken de derdeklassers nogal wat grammaticale en idiomatische fouten: gemiddeld zijn zo'n 30 à 45 procent van de zinnen in dit opzicht inadequaat. Ook worden er bij alle schrijftaken zeer veel interpunctiefouten gemaakt; bij de sollicitatiebrieven bedraagt het gemiddelde zelfs één interpunctiefout per zin!

3 *De ontwikkeling en toepassing van normen*

Het stellen van normen is niet eenvoudig. Er zijn verschillende manieren om normen te ontwikkelen (Meskauskas, 1976; Glass, 1976; Smith & Smith, 1988). Het meest voor de hand liggend is waarschijnlijk het bevragen van relevante groepen respondenten, zoals docenten, didactici, ouders, beleidsvoerders en vertegenwoordigers van het bedrijfsleven. Aan hen wordt dan gevraagd wat de leerlingen ten minste zouden moeten kennen en kunnen. Aan deze methode kleven vooralsnog enkele nadelen. Ten eerste blijken de respondenten nogal te verschillen in hun oordeel over het minimaal wenselijke niveau (Blok, 1987; De Glopper, 1988; Baltzer, De Glopper & Van Schooten, 1988). Wat de één als een voldoende prestatie beschouwt, wordt door de ander als volstrekt inadequaat afgedaan. Op afzonderlijke opgaven kan de discrepantie tussen de beoordelaars zelfs oplopen tot honderd procent (Blok & Van Schooten, 1986). Ten tweede is herhaaldelijk gebleken dat de respondenten hun normen veel te hoog stellen. Zo hoog zelfs dat ze in de praktijk slechts door een handjevol leerlingen gehaald (kunnen) worden (Wijnstra, 1988). Het lijkt alsof men consequent overvraagt. Toepassing van dermate strenge normen betekent dat bij elke opdracht een zeer hoog percentage van de leerlingen beneden de gestelde norm presteert.

Wanneer de normen eenmaal zijn vastgesteld ontstaat een volgend probleem. De scores op de verschillende toetsen moeten namelijk tot één beslissing gecombineerd worden. Dit kan op verschillende manieren. In de praktijk van het voortgezet onderwijs komen veelal compensatoire beslissingen voor; een leerling kan een onvoldoende voor het éne vakonderdeel compenseren door een betere prestatie voor een ander vakonderdeel. Beslissingen worden dan gebaseerd op een somscore of een gemiddelde over alle getoetste leerstofonderdelen. Bij de afsluiting van de basisvorming ligt een dergelijke praktijk niet erg voor de hand. In overeenstemming met het basale en funderende karakter van de basisvorming zal een leerling voor alle eindtermen tenminste het algemene niveau moeten bereiken. In het geval van het vak Nederlands dient zowel voor lezen, schrijven, spreken, taalbeschouwing en informatievaardigheden aan de minimumeisen te worden voldaan. Alle eindtermen zijn immers (bij benadering) van gelijk belang voor vervolgonderwijs, beroep of samenleving. De beslissing zal dan ook conjunct moeten zijn: een leerling kan alleen slagen indien alle individuele beslissingen positief zijn (Wijnen, 1977). Onze gegevens stellen ons echter niet in staat de 'zak/slaag beslissingen' compleet conjunct te maken. In dit artikel maken wij 'zak/slaag beslissingen' voor de onderdelen lezen, schrijven en spreken apart. In feite presenteren wij dus zeer voorzichtige (optimistische) schattingen van de consequenties van het stellen van normen.

3.1 *Een statistische methode*

In de eindtermenontwerpen (Ministerie van Onderwijs en Wetenschappen, 1989) wordt het algemene niveau omschreven als het niveau dat voor tachtig procent van de leerlingen haalbaar geacht wordt; het hogere niveau zou door veertig procent van hen gehaald moeten worden. In feite is hier sprake van een statistisch criterium. Daarbij is de grens tussen een voldoende en onvoldoende prestatie niet zozeer gebaseerd op een inhoudelijke analyse van een opdracht als wel gerelateerd aan hetgeen door een bepaald percentage van de leerlingen bereikt wordt. De norm voor de beheersing van een eindterm, gegeven een betrouwbare en valide toets, is daarmee gemakkelijk vast te stellen door uit te rekenen wat de hoogste toetsscore is die nog net door tachtig (of veertig) procent van de leerlingen bereikt wordt.

Normen voor lezen, spreken en schrijven
In de aan het vakonderdeel lezen gewijde passages uit het eindtermenontwerp Nederlands worden naast eindtermen op het terrein van de deelvaardigheden (het herkennen van het hoofdonderwerp, de hoofdgedachte of de bedoeling van de auteur) ook meer functionele eindtermen voorgesteld, waaronder het leren hanteren van teksten die de lezer tot handelen aansporen. De leerling moet bij de afsluiting van de basisvorming dus aan beide typen doelstellingen van het leesonderwijs te voldoen. Grofweg wil dit zeggen: de leerling moet niet alleen een tekst met begrip kunnen lezen, maar ook een opdracht naar aanleiding van een tekst kunnen uitvoeren. Ter operationalisatie van deze twee elementaire doelen hebben wij voor het onderdeel lezen op basis van het taakboekje 'Schoolreis naar Waddenoog' twee subtoetsen samengesteld. De eerste subtoets indiceert 'functionele leesvaardigheid' en omvat de eerste drie (functionele) leesopdrachten; de tweede subtoets meet regulier tekstbegrip en is opgebouwd uit de vierde en vijfde opdracht van 'Schoolreis naar Waddenoog'. Aangezien deze leesopdrachten ook aan zesdeklassers van het basisonderwijs zijn voorgelegd, kunnen we tevens enig idee krijgen van het percentage basisschoolleerlingen dat aan de eisen van de basisvorming zou voldoen.

Ook de normen voor het vakonderdeel spreken/luisteren zijn gebaseerd op twee indicatoren. De eerste indicator is de totaalscore voor de opdracht 'Randstad biedt aan' en kan worden opgevat als een operationalisatie van de eindterm: *'de leerlingen kunnen van een voor hen bestemde uitleg (...) de belangrijkste elementen vertellen' (p. 22).* De tweede indicator is de totaalscore voor 'Arbeidsvoorwaarden bij Randstad' en vormt een nadere uitwerking van de eindterm: *'de leerlingen kunnen al dan niet telefonisch aan een dialoog deelnemen (...) in een gesprek met een officiële instantie, daartoe kiezen zij geschikte conventies in taalgedrag en taalgebruik en hanteren zij de voor hen juiste strategieën om de nodige informatie te geven of verkrijgen' (p. 21).*

Voor het onderdeel schrijven is van twee taken gebruik gemaakt: de sollicitatiebrief en het persoonlijk getint verhaal over een angstige ervaring dat de leerling voor de schoolkrant moest schrijven. Bij elke taak zijn weer twee indicatoren geconstrueerd. Voor de sollicitatiebrief zijn dat de totaalscore voor de beoordelingen van de algemene aspecten globale kwaliteit, inhoud, stijl en organisatie en de totaalscore van het aantal spelfouten en het aantal werkwoordfouten. Deze eerste indicator kan worden opgevat als een maat voor de algemene kwaliteit van de sollicitatiebrief ofwel voor de mate waarin de leerling 'kan schrijven'; de tweede indicator als een maat voor spelling. Beide indicatoren vormen tezamen en operationalisatie van de eindterm: *'de leerlingen kunnen een zakelijke of persoonlijke brief schrijven (...) ten behoeve van (...) toekomstige werkgevers en daartoe de noodzakelijke conventies hanteren met betrekking tot (...) spelling (...) (p. 23).* Bij het verhaal over een angstige ervaring zijn de beide indicatoren op soortgelijke wijze samengesteld. Hiermee is een operationalisatie verkregen van de eindterm: *'de leerlingen kunnen een persoonlijke impressie beschrijven ten behoeve van hun directe omgeving (school en thuis) en daartoe de noodzakelijke conventies met betrekking tot (...) spelling (...) hanteren (p. 23).* Alle hier genoemde eindtermen betreffen het algemene niveau van de basisvorming Nederlands.

Bevindingen voor lezen, spreken en schrijven
De scores die indicatief zijn voor de niveaus die nog net door 80 en 40 procent van de leerlingen behaald kunnen worden zijn weergegeven in Tabel 3. Omwille van de onderlinge vergelijkbaarheid zijn deze grensscores doorgaans uitgedrukt als een percentage van de maximale score die de leerlingen voor de taak konden behalen. De gemiddelde score voor de algemene kwaliteit van het verhaal is uitgedrukt op een schaal met een gemiddelde van 100 en een standaarddeviatie van 15. Bij het onderdeel spelling is het absolute aantal fouten weergegeven.

Uit Tabel 3 blijkt dat het niveau dat door 80 procent van de leerlingen wordt bereikt voor lezen ligt bij 52 procent van de opgaven voor functioneel tekstbegrip en 48 procent van de opgaven voor regulier tekstbegrip. Dit zou betekenen dat een leerling, wil hij bij dit onderdeel voor het algemene niveau van de basisvorming slagen, aan deze relatief lage scores reeds voldoende zou hebben. Voor het hogere leesniveau bedragen deze grensscores respectievelijk 84 procent (functionele leesvaardig-

Tabel 3 *Het niveau dat door 80 respectievelijk 40 procent van de derdeklassers bereikt werd*

	Algemene niveau (door 80% bereikt)	Hogere niveau (door 40% bereikt)
Lezen:		
functionele leesvaardigheid	51.6	77.1
regulier tekstbegrip	48.4	75.0
Gespreksvaardigheid:		
randstad biedt aan	68.4	84.2
arbeidsvooreaarden bij Randstad	30.0	50.0
Schrijven sollicitatiebrief:		
algemene kwaliteit	12.5	31.3
aantal (werkwoord)spelfouten	11.0	6.0
Schrijven persoonlijk getint verhaal:		
algemene kwaliteit	84.9	95.0
aantal (werkwoord)spelfouten	25.0	12.5

heid) en 50 procent (regulier tekstbegrip). Anders geformuleerd: 40 procent van de leerlingen voldoet aan deze eisen, terwijl ze voor 60 procent te streng zijn.

Om voor het algemene spreekvaardigheidsniveau te kunnen slagen hoeft de leerling slechts 68 procent van de opgaven van 'Randstad biedt aan' en hooguit 30 procent van de opgaven van 'Arbeidsvoorwaarden bij Randstad' correct te hebben beantwoord; voor het hogere niveau, dat voor 40 procent van de leerlingen haalbaar is, liggen de normen bij resp. 84 en 50 procent van de opgaven.

Met betrekking tot de eindtermen 'het schrijven van een sollicitatiebrief' en 'het schrijven van een persoonlijke impressie...' valt op dat het algemene niveau als uiterst laag kan worden getypeerd. Bij de sollicitatiebrief ligt het niveau dat door 80 procent van de leerlingen gehaald wordt bij één achtste deel van de maximale score voor de algemene kwaliteit, terwijl voor het slagen op het hogere niveau een score van één derde nog toereikend is. Voor het onderdeel spelling zou de leerling voor het algemene niveau maar liefst 11 (sollicitatiebrief) en 25 fouten (verhaal) in de (werkwoord)spelling mogen maken; voor het hogere niveau zijn deze aantallen nog steeds aanzienlijk: respectievelijk 6 en 12.5 fouten. Deze toetsscores geven een indicatie van de prestaties waartoe een gemiddelde derdeklasser in staat moet zijn om voor het algemene c.q. hogere niveau van de basisvorming te kunnen slagen. De 'gemiddelde' derdeklasser is echter een in werkelijkheid niet bestaande abstractie. In de praktijk van het voortgezet

Tabel 4 *Percentage leerlingen per schooltype dat niet aan de norm voor de basisvorming Nederlands zou voldoen*

	VWO	HAVO	MAVO	LTO	LHNO	POPULATIE
Algemene niveau:						
lezen	3	4	13	49	48	20
spreken	4	4	19	36	45	20
sollicitatiebrief	3	10	17	52	30	20
verhaal	4	8	12	51	33	20
Hogere niveau:						
lezen	26	31	61	89	90	60
spreken	29	37	66	80	82	60
sollicitatiebrief	32	45	60	87	71	60
verhaal	29	40	60	90	84	60

onderwijs zijn leerlingen verdeeld over diverse schooltypen. Uit de Proefpeiling Nederlands komt naar voren dat deze schooltypen qua gemiddeld prestatieniveau aanzienlijk van elkaar verschillen (Kuhlemeier & Van den Bergh, 1989b). Het is dan ook niet van enig onderwijskundig en maatschappelijk belang ontbloot om meer te weten over de samenstelling van de groep leerlingen die niet aan onze criteria voor de basisvorming Nederlands zou voldoen. In Tabel 4 zijn enkele gegevens vermeld betreffende de samenstelling naar schooltype.

Uit Tabel 4 blijkt dat bij toepassing van de 80-procentsnorm op de PPON-taken en de huidige groep derdeklassers, met name op het LBO (LHNO plus LTO) kan leiden tot grote zakpercentages. Dit percentage varieert afhankelijk van het onderdeel van 30 tot 52 procent. In het VWO/HAVO presteert, afhankelijk van het onderdeel, drie à tien procent van de leerlingen onder het minimumniveau. Voor het hogere niveau zijn de uitkomsten in onze ogen nog verontrustender. In het LBO weet slechts een enkeling het hogere niveau te bereiken. Het percentage 'beneden-minimale' leerlingen varieert van taak tot taak van 71 tot 90 procent. Zij scoren lager dan hetgeen door veertig procent van de totale groep derdeklassers gepresteerd wordt. Het percentage VWO/HAVO-leerlingen dat voor het hogere niveau zou 'zakken', varieert al naar gelang de taak van een kwart tot bijna de helft.

3.2 *Inhoudelijk beargumenteerde normen*

Een tweede methode om normen te stellen, de kernitemmethode (De Groot & Van Naerssen, 1969), is gebaseerd op een inhoudelijke vertaling van de minimumeisen in toetstermen. Daarbij wordt een aantal opgaven aangewezen die een kandidaat tenminste correct moet hebben om voor een proefwerk of examen te kunnen slagen. Deze opgaven markeren de ondergrens van beheersing. Bij onze keuze van de kernopgaven staat de vraag centraal: kan de communicatie op basis van de geleverde prestatie slagen? Voor een aantal opdrachten hebben wij bepaald welke elementen vanuit communicatief oogpunt essentieel zijn. Leerlingen die niet aan deze minimale communicatie-eisen voldoen, leveren, naar onze uiteraard subjectieve mening, een onvoldoende prestatie. Deze methode is in ons land voor het eerst toegepast door het onderzoeksteam dat de Voorstudie Periodieke Peiling van het Onderwijsniveau in het basisonderwijs uitvoerde (Wesdorp e.a., 1985; Van den Bergh, 1988). Waar mogelijk en wenselijk hebben wij in de Proefpeiling Nederlands een soortgelijke methode toegepast (Kuhlemeier & Van den Bergh, 1989b). We illustreren deze methode aan de hand van de drie spreektaken: 'Randstad biedt aan', 'Arbeidsvoorwaarden bij Randstad Uitzendbureau' en 'Het navertellen van een verhaal'. Voor de resterende vakonderdelen worden alleen de resultaten gepresenteerd.

Normen voor de spreektaken

In de beide telefoonopdrachten 'Randstad Uitzendbureau' en 'Arbeidsvoorwaarden bij Randstad' moet de leerling informatie inwinnen over een aangeboden baan. Voor het eerste telefoongesprek kan een communicatief doel gesteld worden: het verkrijgen van een vakantiebaantje. Om het aangeboden baantje te krijgen moet de leerling zich (kunnen) melden voor een kennismakingsgesprek met de bedrijfsleider. Concreet betekent dit dat ten minste de naam van het bedrijf, de vestigingsplaats, het adres, de datum en tijdstip van het gesprek alsmede de route die gevolgd moet worden om bij het bedrijf te komen onthouden of genoteerd moeten zijn. Leerlingen die één of meer vragen naar deze elementen niet weten te beantwoorden, kunnen niet naar het kennismakingsgesprek; het doel van het telefoongesprek, het verkrijgen van vakantiewerk, wordt niet bereikt. Van het tweede telefoongesprek met Randstad lijken twee zaken het belangrijkst: de studiekostenvergoeding en de kinderbijslag. De leerling moet, in onze ogen, weten of de kinderbijslag en/of de studiekosten in 'gevaar' komt door het accepteren van dit vakantiebaantje. Als de prestaties op deze opdracht stellen wij derhalve de eis dat tenminste de drie vragen over deze beide onderwerpen correct beantwoord zijn. Voor de derde spreekopdracht, over de produktie van bakstenen, is het zeer lastig om inhoudelijk te beargumenteren welke prestaties als minimaal en welke prestaties als voldoende mogen worden beschouwd. Wat betreft de inhoud van de explicatie van de leerling hebben wij de eis gesteld dat de leerling tenminste zeven van de negentien inhoudselementen correct moet noemen. Ten aanzien van de wijze waarop de leerling de uitleg

presenteert (taalgebruik en opbouw) is de grens gesteld op 2 (uit maximaal 7). Alleen leerlingen die zowel voor de inhoud als voor de presentatiewijze beneden genoemde grensscores presteren, leveren in onze ogen een minimale spreekprestatie.

In Tabel 5 zijn de consequenties van dergelijke inhoudelijke normen voor spreken en lezen weergegeven. In deze tabel komt tot uitdrukking welk percentage van de derdeklassers op geen, één, twee of meer opdrachten beneden het minimale niveau presteerde. Voor schrijven hebben we een dichotomie gemaakt: beide opdrachten beneden de grensscore of ten minste één opdracht boven de grensscore.

Tabel 5 *Percentages leerlingen die niet aan de gestelde normen voldoen*

Aantal taken waarop beneden minimaal niveau gepresteerd is	Spreken	Lezen	Schrijven
0	58	47	68
1	33	28	nvt*
2	8	13	32
3	1	8	
4	nvt	3	
5		1	

* Voor schrijven is een dichotomie gemaakt: beide opdrachten minimaal tegenover niet beide opdrachten minimaal.

Uit Tabel 5 blijkt dat het stellen van inhoudelijke beargumenteerde normen grote gevolgen kan hebben, te meer wanneer de zak/slaag beslissing conjunct genomen wordt. Hoewel wij getracht hebben per taak zeer soepele normen te formuleren is het resultaat (in onze ogen) zorgwekkend. Zo varieert voor de verschillende leesopdrachten het percentage leerlingen met een score lager dan de gestelde norm van zeven tot dertig. Echter, slechts 47 procent heeft een boven-minimale score voor alle vijf de leesopdrachten. Een gecombineerde eis voor alle opdrachten lijkt 'dodelijk'.

Ook voor het relatief 'makkelijke' onderdeel spreken slaagt op basis van onze normen slechts een kleine meerderheid. Voor het onderdeel schrijven lijken de consequenties van een inhoudelijk beargumenteerde norm desastreus; slechts 18 procent van de leerlingen voldoet aan een inhoudelijke norm. Misschien is onze norm hier te streng geweest. Echter, voor bij voorbeeld 'Het excuusbriefje' kon een leerling volstaan met vijf inhoudselementen (de naam van de leerling, refereren aan de gemaakte afspraak met de directeur en het tijdstip van de afspraak, de reden van verhindering aangeven en in verband met het formele karakter van het briefje een verontschuldiging formuleren). Merk op dat er geen eisen gesteld zijn aan de toon, de organisatie of formele aspecten (spelfouten, grammaticale fouten e.d.). Voor de sollicitatiebrief geldt overigens hetzelfde; ook hier zijn eisen geformuleerd die alleen betrekking hebben op de aanwezigheid van bepaalde inhoudselementen.

Het lijkt niet voorbarig te stellen dat een inhoudelijke, op het slagen van de communicatie gebaseerde norm leidt tot gigantische zakpercentages. In dit verband geven de hier gepresenteerde gegevens nog een veel te rooskleurig beeld. Immers, de gegevens zijn voor de drie aspecten *apart* geanalyseerd. Terwijl de eindtermen een duidelijk conjunct karakter hebben. Een 'educated guess' is dat bij een conjuncte beslissing voor alle opdrachten géén van de leerlingen geslaagd zou zijn.

3.3 *Een inhoudelijk-statistische methode*

In het eindtermenontwerp wordt naast de voorheen besproken statistische methode ook een meer inhoudelijk beargumenteerd voorstel gedaan. Daarbij is er sprake van een hoger niveau naarmate in een leerlingprestatie meer onderwerpen op een meer gedetailleerde wijze naar voren komen en naarmate een hogere mate van abstractie en complexiteit wordt bereikt (Eindtermencommissie Nederlands, 1989). In het vervolg van deze paragraaf wordt dit laatste criterium voor twee taken uit de Proefpeiling Nederlands nader uitgewerkt. Bovendien wordt, gegeven de gestelde criteria, de vraag naar de haalbaarheid van de eindtermen aan de hand van de feitelijke leerlingprestaties nader geëxploreerd. Gekozen is voor de spreektaak 'Het navertellen van een verhaal' en de schrijftaak 'Een sollicitatiebrief'. De spreektaak kan worden opgevat als een operationalisatie van de eindterm: *'De leerlingen kunnen in de vorm van een korte monoloog voor een hun bekend publiek een persoonlijke beleving vertellen of een situatie beschrijven' (p. 22).* De taak 'Een sollicitatiebrief' vormt een operationalisatie van de eind-

term: *'De leerlingen kunnen een zakelijke of persoonlijke brief schrijven (. . .) ten behoeve van (. . .) toekomstige werkgevers en daartoe de noodzakelijke conventies hanteren met betrekking tot interpunctie, spelling, structuur, kenmerken van tekstsoorten, taalgebruik en uiterlijke verzorging (. . .)' (p. 23).* Beide hier geciteerde eindtermen betreffen het algemene niveau van de basisvorming. Volgens het eindtermenontwerp zou na de afsluiting van de basisvorming dus tachtig procent van de leerlingen tot een dergelijke prestatie in staat moeten zijn.

Het navertellen van een verhaal

Bij de spreektaak 'Het navertellen van een verhaal' onderscheiden we twee aspecten: de inhoud van het verhaal en de wijze waarop het verhaal naar voren wordt gebracht (taalgebruik en opbouw). Uit Tabel 1 valt op te maken dat de derdeklassers gemiddeld de helft van de inhoudselementen correct wisten te noemen, terwijl het gemiddeld voor taalgebruik en opbouw 63 procent van de maximaal te behalen score bedroeg. Deze gegevens bieden weliswaar enig inzicht in de spreekprestaties van de leerlingen. Over de vraag hoeveel leerlingen de eindterm *'de leerlingen kunnen in de vorm van een korte monoloog voor een hun bekend publiek een situatie beschrijven'* daadwerkelijk bereiken, moet eerst nog een criterium worden bepaald dat aangeeft welke prestaties als adequaat en welke als inadequaat mogen worden beschouwd. In onze exploratie van de niveauproblematiek kiezen wij voor een deels inhoudelijke-en deels statistische benadering. In het geval van 'Het navertellen van een verhaal' kan een leerling voor 0, 1 of 2 aspecten een (on)voldoende prestatie behalen. Naar onze mening dient een leerling zowel voor het aspect inhoud als voor taalgebruik/organisatie een adequate prestatie te leveren. Een goed gepresenteerd, maar inhoudelijk armoedig verhaal is immers evenzeer inefficiënt als een inhoudelijk adequaat, maar slecht gepresenteerd verhaal. Omdat een inhoudelijke verantwoording van het onderscheid tussen voldoende en onvoldoende prestatie voor de beide aspecten moeilijk te geven is, wordt de mate van beheersing uitgedrukt als een percentage van de maximaal te behalen score voor het desbetreffende aspect. Daarbij onderscheiden we de normen 20, 40, 60 en 80 procent. Voor het aspect inhoud komt een norm van 20 procent bij voorbeeld overeen met vier correct weergegeven inhoudselementen; bij het aspect taalgebruik/organisatie komt deze norm overeen met een score van 2 (op een zevenpuntschaal). In Tabel 6 is voor vier van de zestien mogelijke combinaties van beide criteria het percentage leerlingen dat een adequate prestatie levert weergegeven.

Uit Tabel 6 blijkt dat het meest soepele criterium door 94 procent van de leerlingen bereikt wordt; zes procent van de leerlingen voldoet niet aan deze norm. Deze leerlingen weten hooguit vier van de negentien inhoudselementen correct te noemen, terwijl hun taalgebruik en de organisatie van het verhaal in hoge mate te wensen overlaat (score 1 of 2). Het meest strenge criterium wordt daarentegen slechts door twee procent van de derdeklassers gehaald. Zij noemen tenminste vijftien inhoudselementen, terwijl hun score voor taalgebruik/organisatie 6 of 7 bedraagt[2]. Er zou een keuze gemaakt kunnen worden voor een spreekprestatie die men nog net voldoende acht. Misschien vinden sommigen deze op 'niveau 2' liggen, anderen zullen een keuze maken voor 'niveau 3' en weer anderen ergens tussen niveau 1 en 2 in. Bij een dergelijke keu-

Tabel 6 *Percentage leerlingen dat voor beide aspecten van 'Het navertellen van een verhaal' tenminste 20, 40, 60 of 80 procent van de maximale score behaalt en het percentage leerlingen dat voor beide aspecten van 'De sollicitatiebrief' aan de gestelde eisen voldoet*

'Het navertellen van een verhaal'

Score inhoud:	Score taalgebr. en org. 20%	40%	60%	80%
20% correct	94%			
40% correct		60%		
60% correct			20%	
80% correct				2%

'De sollicitatiebrief'

Aantal formele fouten	Score voor 'het schrijven' 20%	40%	60%	80%
3	25%			
2		11%		
1			2%	
0				0.1%

Tabel 7 *De percentages leerlingen met een score onder het criterium voor 'Het navertellen van een verhaal' en 'De sollicitatiebrief'*

Schooltype	Opdracht	Niveau 1	2	3	4
LHNO	verhaal navertellen	12	67	94	100
LTO	verhaal navertellen	4	52	81	100
MAVO	verhaal navertellen	4	42	80	100
HAVO	verhaal navertellen	4	30	76	97
VWO	verhaal navertellen	2	19	62	92
LHNO	sollicitatiebrief	87	87	99	100
LTO	sollicitatiebrief	94	99	100	100
MAVO	sollicitatiebrief	76	89	97	100
HAVO	sollicitatiebrief	69	87	98	100
VWO	sollicitatiebrief	53	75	85	99.9

ze, die altijd een enigszins arbitrair karakter zal dragen, moet niet alleen inhoudelijk gefundeerd worden (op basis van concrete leerlingprestaties), maar er moet ook rekening gehouden worden met het percentage leerlingen dat niet aan de gestelde eis voldoet. Anders loopt men licht de kans normen te stellen die weliswaar wenselijk zijn, maar die in de praktijk voor velen niet haalbaar zijn. Vandaar dat in Tabel 7 per 'niveau' een overzicht gegeven wordt van het percentage leerlingen per schooltype dat een score beneden de gestelde normen behaalt.

Uit Tabel 7 wordt duidelijk dat er tussen de vijf schooltypen een grote variatie is in het percentage leerlingen wiens prestatie niet aan de norm voldoet. Bij de meest coulante norm voor deze spreekopdracht voldoet 18 procent van de prestaties van de LHNO-leerlingen niet; voor VWO-leerlingen bedraagt dit percentage twee procent. Ook wordt duidelijk dat de combinatie van twee eisen (voor: 'inhoud' en 'taalgebruik en organisatie') kan leiden tot een groot aantal leerlingen met een ontoereikende prestatie. In dit verband wijzen wij op de toch aanzienlijke percentages leerlingen met onvoldoende prestaties indien 'niveau 2' aangehouden zou worden. In onze ogen zijn deze percentages zo hoog dat dit niveau in onderwijskundig opzicht als onhaalbaar voor veel leerlingen en in maatschappelijk opzicht als onaanvaardbaar gekarakteriseerd moet worden.

De sollicitatiebrief

In het geval van een sollicitatiebrief zijn ten minste twee aspecten van belang. De brief moet niet alleen 'goed geschreven' zijn, maar mag bovendien geen formele taalfouten bevatten. Het eerste aspect is geoperationaliseerd als de algemene kwaliteit van de brieven. Het tweede aspect is geoperationaliseerd als de som van de absolute aantallen spelfouten, werkwoordspelfouten en grammaticale en idiomatische fouten. Opgemerkt zij dat het aantal interpunctiefouten bij dis aspect geen rol speelt. Met betrekking tot deze formele vereisten stellen we de grens op 0, 1, 2 of 3 taalfouten. Ten aanzien van het algemene kwaliteitsoordeel over de sollicitatiebrieven stellen we de grens op 20, 40, 60 of 80 procent van de maximale score. In Tabel 6 is voor vier van de zestien mogelijke combinaties van beide criteria het percentage leerlingen dat een adequate prestatie levert weergegeven.

Tabel 6 wijst uit dat het meest soepele criterium door een kwart van de leerlingen bereikt wordt; drie kwart van hen voldoet niet aan deze norm. Zij maken meer dan drie taalfouten (spelfouten, werkwoordspelfouten en/of grammaticale of idiomatische fouten), terwijl de inhoud, de stijl, de organisatie en de algemene kwaliteit van hun briefjes in hoge mate te wensen overlaat (score hooguit twintig procent van de maximale score). Aan het meest strenge criterium voldoet slechts één van de ruim 1400 leerlingen uit de responsgroep. Hij of zij zondigt in het briefje niet één keer tegen de taalconventies op het gebied van spelling en grammaticaliteit, terwijl de score voor 'het schrijven' tenminste 80 procent van de maximale score bedraagt[2].

Een indicatie van het niveau van de 'gemiddelde' derdeklasser heeft zonder referentie

aan het schooltype weinig zeggingskracht (zie Tabel 7). Uit Tabel 7 blijkt dat het laagste van de vier onderscheiden niveaus door 6 procent van de LTO-leerlingen en door 47 procent van de VWO-leerlingen wordt gehaald. Aan de strengere eisen op het vierde niveau voldoet vrijwel geen van de sollicitatiebrieven. Kortom, zelfs bij de meest soepele eis presteert nog een groot deel van de leerlingen onder de norm. Toch kan deze eis niet zwaar genoemd worden. Eigenlijk mag in een sollicitatiebriefje geen enkele formele fout staan. Bovendien speelt de interpunctie geen rol in het criterium, anders zou het resultaat nog zorgwekkender zijn. Bij het meest strenge criterium, wat overigens een zeer goede sollicitatiebrief opleverde, behaalde slechts één VWO-leerling een prestatie boven de norm.

Wanneer de feitelijke produkten in ogenschouw genomen worden, dan zullen sommigen er (terecht) op wijzen dat het sollicitatiebriefje bij niveau 1 zo slecht nog niet is; wij zijn het daar in elk geval van harte mee eens. Maar let wel: dit briefje behoort tot de bovenste 25 procent van de leerlingprodukten.

4 *Slotbeschouwingen*

In onze verkenning van de normstellingsproblematiek hebben wij drie manieren om normen te ontwikkelen uitgeprobeerd. Voor welke manier men ook kiest, telkens leiden onze explorations tot uitkomsten die wij als zorgwekkend typeren. Het 'communale', door 80 procent van de leerlingen bereikte niveau, blijkt inhoudelijk gezien zeer mager te zijn, terwijl nogal wat leerlingen voor de schrijf-, lees-, spreek- en luistertaken een minimale of twijfelachtige prestatie leveren. Grote groepen (LBO-)leerlingen zouden zelfs voor het algemene niveau zakken. Deze uitkomsten moeten worden geïnterpreteerd in het licht van de wijze waarop ze zijn verkregen, vandaar dat enkele kanttekeningen op hun plaats zijn.

Een eerste kanttekening betreft de wijze waarop de normen ontwikkeld zijn. Elk van de drie gehanteerde methoden heeft zijn eigen voor- en nadelen. Eén van de voordelen van de aan het eindtermenontwerp ontleende statistische methode is de eenvoud en de toetsbaarheid. De normen voor het algemene en het hogere niveau zijn, zoals we hebben laten zien, eenvoudig vast te stellen door uit te rekenen wat de toetsscores zijn die nog net door 80 respectievelijk 40 procent van de leerlingen bereikt worden. Een zuiver statistische methode draagt echter het gevaar van willekeur in zich. De keuze voor een bepaald percentage leerlingen is immers tamelijk arbitrair en zonder nadere inhoudelijke argumentatie niet goed te verantwoorden. Een methode waarbij de normen via een inhoudelijke analyse van de eindtermen en de daarbij behorende toetsen worden verkregen verdient daarom in principe de voorkeur. Over de wijze waarop wij deze inhoudelijke methode, de kernitemmethode, hebben toegepast is echter discussie mogelijk. Met name resteert twijfel aan de objectiviteit, de betrouwbaarheid en de validiteit van *onze* op het kunnen slagen van de communicatie gebaseerde normen. Dat dit een subjectieve en daarmee minder wenselijke gang van zaken is, behoeft geen betoog. Bovendien heeft de kernitemmethode tot nadeel dat beslissingen genomen worden op een gedeelte van de informatie (alleen de scores op de essentiële items worden in de analyse betrokken).

Ook bestaat de mogelijkheid dat onze operationalisaties de desbetreffende eindtermen onvolledig of op anderszins inadequate wijze representeren. Ook zouden onze toetsen in relatie tot de beoogde eindtermen wellicht te moeilijk of juist te gemakkelijk kunnen zijn geweest. De gerapporteerde percentages minimale presteerders dienen derhalve met gepaste terughoudendheid te worden geïnterpreteerd. Wel dient opgemerkt te worden dat wij uiterst voorzichtig zijn geweest in het kiezen van de grensscores. Dat wil zeggen: ook in communicatief opzicht zeer magere leerlingprestaties zijn nog geaccepteerd als 'bewijs' van kunnen schrijven, lezen, spreken of luisteren.

De derde methode lijkt voor de toekomst nog het meeste perspectief te bieden, aangezien deze de voordelen van zowel de statistische als de inhoudelijke methode in zich verenigt en de diverse onderscheiden niveaus aan de hand van concrete leerlingprodukten verhelderd worden. Hiermee wordt een betere basis gelegd voor een beredeneerde keuze voor een bepaald niveau als minimumdoel om naar te streven. Het verdient aanbeveling deze methode nader uit te werken, zo mogelijk in samenwerking met degenen die betrokken zijn bij de ontwikkeling van eindtermen en de constructie van toetsen ter afsluiting van de basisvorming.

Een tweede kanttekening betreft onze interpretatie van de basisvorming en onze keuze en operationalisatie van concrete eindtermen uit het eindtermenontwerp Nederlands. In dit ontwerp worden vele concrete eindtermen voor het algemene niveau beschreven. Die beschrijven zoals gezegd hetgeen (vrijwel) alle leerlingen aan het einde van de basisvorming ten minste zouden moeten kennen en kunnen. Het basale en fundamentele karakter van de basisvorming geeft aanleiding tot de veronderstelling dat een leerling al deze eindtermen moet beheersen om het vak Nederlands op het algemene niveau te kunnen afsluiten. Voorts worden in het eindtermenontwerp Nederlands bij iedere eindterm meer kwaliteitsaspecten genoemd. Bij voorbeeld: *'De leerlingen kunnen een zakelijke of persoonlijke brief schrijven (...) ten behoeve van (...) toekomstige werkgevers en daartoe de noodzakelijke conventies hanteren met betrekking tot interpunctie, spelling, structuur, kenmerken van tekstsoorten, taalgebruik en uiterlijke verzorging' (p. 23).* Ook hier zijn wij ervan uitgegaan dat voor een beheersing van een eindterm aan alle genoemde aspecten moet worden voldaan. Het gaat immers om het minimaal noodzakelijke. Hierdoor geïnspireerd hebben wij in onze operationalisatie van eindtermen telkens twee aspecten geoperationaliseerd. Een sollicitatiebrief moet bij voorbeeld niet alleen goed geschreven zijn, maar mag ook niet al te veel spelfouten bevatten. De keuze voor twee aspecten is tot op zekere hoogte arbitrair. Met evenveel recht hadden wij bij voorbeeld ook het aantal interpunctiefouten of de netheid van het handschrift in de norm kunnen opnemen. In onze operationalisatie van een eindterm moet een leerling aan beide geoperationaliseerde aspecten *tegelijk* voldoen. Indien we de leerlingprestaties aan deze norm spiegelen, dan zijn de resultaten in onze ogen desastreus: grote groepen leerlingen blijken zelfs beneden relatief soepele normen te presteren. Bovendien blijkt het hogere niveau bij alle door ons gekozen eindtermen voor vrijwel alle LBO-leerlingen te hoog gegrepen. En dan hebben we het daarbij pas over twee aspecten van telkens één van de vele doelstellingen uit het eindtermenontwerp Nederlands! Hadden we de eis gesteld dat een leerling voor meer eindtermen van het vak Nederlands op het algemene niveau moet presteren, dan hadden de uitkomsten nog meer aanleiding tot zorg gegeven. Wanneer we daarnaast ook nog bedenken dat de leerlingen voor de toelating tot HAVO/MBO behalve Nederlands nog zes andere vakken op het algemene niveau moeten afsluiten, dan moeten we haast wel aannemen dat de HAVO/MBO-docenten een zeer rustige tijd te wachten staat. Het verdient derhalve aanbeveling de thans voorliggende eindtermen nog eens kritisch te beschouwen; wellicht wordt hier en daar toch overvraagd. Misschien zou er bij het formuleren van eindtermen meer rekening gehouden kunnen worden met de feitelijke prestaties van leerlingen zoals deze uit peilingsonderzoek naar voren komen. Wij zouden gezien de huidige eindtermen de stelling willen verdedigen dat: reële eindtermen, waarbij 80 procent van de leerlingen kan slagen, enkel geformuleerd kunnen worden op het niveau: 'Zeg het alfabet op'; het hogere niveau wordt gekenmerkt door het alfabet achterstevoren op te zeggen, waarbij de leerling zich tenminste vijfmaal mag vergissen.

Een op de vorige aansluitende kanttekening betreft de mogelijkheid en de zinvolheid van het formuleren van één minimumniveau voor allen. In hoeverre kunnen eindtermen die een minimumniveau voor allen beschrijven nog functioneren als doelen waar een ieder naar zou moeten streven? Eén manier om een groter deel van de leerlingen te doen slagen voor de basisvorming is het verlagen van de normen. Zijn er niet veel te veel eindtermen en aspecten geformuleerd? Moeten de leerlingen niet te veel leerinhouden kennen? Wordt niet een te hoge beheersingsgraad geëist? Zou men de eis van zeven of negen voor de doorstroming vereiste vakken niet wat kunnen afzwakken? Uit onze exploratie komt naar voren dat het door tachtig procent van de leerlingen bereikte niveau als niet erg hoog kan worden gekarakteriseerd. De zinvolheid van het verlagen van de normen moet ons inziens dan ook sterk worden betwijfeld. Gevreesd moet worden dat de eindtermen en de normen dan een triviaal en betekenisloos karakter zullen krijgen. Het verwachte positieve effect op het niveau van het onderwijs en de leerlingprestaties zal dan per definitie een zelfde omvang hebben: bijkans minimaal. Gezien de grote prestatieverschillen tussen leerlingen moet bovendien worden gevreesd dat vele van de werkelijk communale eindtermen voor grote groepen leerlingen nauwelijks een uitdaging

vormen. En wel omdat ze deze doelen reeds (denken te) beheersen. Deze verwachting wordt mede ondersteund door het gegeven dat bijna 65 procent van de zesdeklassers uit het basisonderwijs voor het vakonderdeel lezen aan het (door ons geformuleerde) algemene niveau bleek te voldoen. Wel moeten we hierbij bedenken dat Nederlands een relatief 'curriculumonafhankelijk' vak is en de bevindingen dus niet zonder meer naar andere vakgebieden gegeneraliseerd mogen worden.

Een vierde kanttekening is dat onze analyse verricht is in een situatie waarin de basisvorming nog niet is ingevoerd. De dataverzameling vond immers plaats in het voorjaar van 1988. De door de eindtermencommissies beoogde veranderingen in het onderwijsaanbod hebben nog geen ingang gevonden, er heeft nog geen additieve leergangontwikkeling plaatsgevonden, de voorziene nascholingsoperaties moeten nog starten, de eindtermen zijn nog in revisie, de toetsen ter afsluiting van de basisvorming zijn nog in ontwikkeling, een deel van de leerlingen mag één jaar langer over de basisvorming doen, in de basisvorming zullen de prestatieverschillen tussen leerlingen wellicht minder groot zijn, et cetera. Van deze beperkingen zijn wij ons wel bewust. Hoewel men geen overdreven verwachtingen over de veranderbaarheid van onderwijs mag koesteren lijkt een zekere niveauverhoging onder een aantal voorwaarden wel degelijk mogelijk. Met andere woorden: onze uitkomsten beschrijven waarschijnlijk de ondergrens van het niveau dat in de basisvorming haalbaar is. Deze relativeringen zijn ons inziens echter onvoldoende reden om de hier gepresenteerde situatie-analyse zonder meer van tafel te vegen. Het is in ieder geval wel de uitgangssituatie, van waaruit de invoering van de basisvorming gestalte moet krijgen.

Tot slot, gegeven de voorheen genoemde beperkingen, mogen op grond van de hier gepresenteerde analyse geen conclusies getrokken worden over de hoogte van het huidige onderwijsniveau. Eerst zal er een oplossing gevonden moeten worden voor de problematiek van het ontwikkelen van algemeen geaccepteerde, objectieve, betrouwbare, valide en realistische normen. Het verdient daarbij aanbeveling om meer rekening te houden met gegevens over de feitelijke leerprestaties van leerlingen zoals die in peilingsonderzoek verzameld worden.

Noten

1. Eigenlijk bevatte dit taakboekje zes opdrachten. Echter, om in de discussie een vergelijking met de prestaties van zesdeklassers uit het basisonderwijs mogelijk te maken, zijn de prestaties op één van de taken niet in dit artikel opgenomen.
2. Om redenen van ruimte zijn geen transcripten van leerlingprestaties opgenomen. Desgewenst kunnen deze bij de auteurs opgevraagd worden.

Literatuur

ARVO, *Advies over eindtermen basisvorming*. Zeist: Adviesraad voor het voortgezet onderwijs, 1989.

Baltzer, J. E., K. de Glopper & E. van Schooten, *De taalvaardigheid van eerstejaars HBO-studenten*. Amsterdam: SCO, 1988.

Bergh, H. van den, Functionele doelstellingen voor het moedertaalonderwijs. *Pedagogische Studiën*, 1988, *65*, 139-148.

Bergh, H. van den & J. B. Kuhlemeier, *De inhoud en het niveau van het vak Nederlands. Periodieke niveaupeiling in het voortgezet onderwijs: een onderzoek naar de haalbaarheid*. 's-Gravenhage: Staatsdrukkerij, 1989.

Blok, H., *Taal voor alledag: Feiten en meningen over het taalgebruik van LBO- en MAVO-leerlingen in alledaagse situaties*. 's-Gravenhage: SVO, 1987.

Blok, H. & E. van Schooten, *Feiten en voorspellingen over het taalgebruik van eindleerplichtige LBO- en MAVO-leerlingen*. Amsterdam: SCO, 1986.

Brinke, J. S. ten, De eindtermen Nederlands voor de basisvorming. Enige toelichtende kanttekeningen. *Levende Talen*, 1989, nr. 441, 316-322.

Eindtermencommissie Nederlands, *Advies over de voorlopige eindtermen basisvorming in het voortgezet onderwijs: Nederlands*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1989.

Glass, G. V., Standards and criteria. *Journal of Educational Measurement*, 1978, *5*, 237-261.

Glopper, K. de, *Schrijven beschreven*. 's-Gravenhage: SVO, 1988.

Groot, A. D. & R. F. van Naerssen, *Studietoetsen: construeren, afnemen en analyseren (deel II)*. Den Haag: Mouton, 1969.

ISCB, *Advies over de eindtermen basisonderwijs en de eindtermen basisvorming voortgezet onderwijs*. Zoetermeer: Inspectie van het Onderwijs, 1989.

Kuhlemeier, J. B. & H. van den Bergh, *Advies betreffende de opzet van een periodiek peilingsonderzoek in het voortgezet onderwijs*. Arnhem: Cito, 1989a.

Kuhlemeier, J. B. & H. van den Bergh, *De Proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs*. Arnhem: Cito, 1989b.

Meskauskas, J. A., Evaluation models for criterion referenced testing: views regarding mastery and standard-setting. *Review of Educational Research*, 1976, *46*, 133-158.

Ministerie van Onderwijs en Wetenschappen, *De ontwikkeling van eindtermen voor het basisonderwijs en de basisvorming in het voortgezet onderwijs (deel 1)*. Zoetermeer: Ministerie van Onderwisj en Wetenschappen, 1989.

PvdA, *Een perspectief voor basisvorming*. Den Haag: Tweede Kamerfractie Partij van de Arbeid, 1989.

Rijlaarsdam, G., *Basisvorming Nederlands*. Leiden: Martinus Nijhoff, 1989a.

Rijlaarsdam, G., Eindtermen basisvorming en evaluatie van het talenonderwijs. *Levende Talen*, 1989b, nr. 441, 340-344.

Smith, R. L. & J. K. Smith, J. K. Differential use of item information by judges using Angoff and Nedelsky procedures. *Journal of Educational Measurement*, 1988, *25*, 259-274.

Wesdorp, H., H. van den Bergh, D. J. Bos, J. B. Hoeksma, R. J. Oostdam, J. Scheerens & B. Triesscheijn, *De haalbaarheid van periodiek peilingsonderzoek. Een voorstudie op het gebied van taalonderwijs in de lagere school*. Lisse: Swets & Zeitlinger, 1985.

Wijgh, I. F. Het eind van de basisvorming mvt. *Levende Talen*, 1989a, nr. 441, 335-339.

Wijgh, I. F., *Frans, Duits en Engels op twee niveaus. Moderne vreemde talen in de basisvorming*. Algemene Publikatie Nr. 62. Arnhem: Cito, 1989b.

Wijnen, W. H. F., Toetsen en beslissen. *Losbladig Onderwijskundig Lexicon*, 1977, *2*, MK 1600, 1-14. Alphen aan de Rijn: Samson.

Wijnstra, J., *Balans van het rekenonderwijs in de basisschool: Uitkomsten van de eerste rekenpeiling medio en einde basisonderwijs*. Arnhem: Cito, 1988.

Curricula vitae

H. van den Bergh (1957) studeerde psychologie aan de Universiteit van Amsterdam. Na zijn studie werkte hij mee aan diverse onderzoeken, o.a. peilingsonderzoek (in het basisonderwijs, speciaal onderwijs en voortgezet onderwijs) en onderzoek naar de validiteit van veranderde examens in Nederland voor het LBO en MAVO. Op dit laatste onderwerp promoveerde hij in 1989.

Adres: Rijksuniversiteit te Utrecht, Afdeling Taalbeheersing, Trans 10, 3512 JK Utrecht.

H. Kuhlemeier (1952) studeerde onderwijskunde aan de Katholieke Universiteit Nijmegen. Hierna was hij als wetenschappelijk medewerker verbonden aan het Sociologisch Instituut te Nijmegen. Thans is hij als onderwijskundige werkzaam bij het CITO te Arnhem, waar hij werkzaamheden op het gebied van de constructie van toetsen voor leesvaardigheid en woordenschat, attitudemeting en leerplanevaluatie verrichtte. Sinds twee jaar is hij leider van het project Periodieke Peiling van het Onderwijsniveau in het Voortgezet Onderwijs (PPON).

Adres: CITO, Afdeling Beginfase Voortgezet onderwijs, Postbus 1034, 6801 MG Arnhem

Manuscript aanvaard 2-10-'89

Summary

Bergh, H. van den & H. Kuhlemeier. 'The realizability of standards Dutch in secondary education'. *Pedagogische Studiën*, 1990, *67*, 1-15

In Dutch politics standards are proposed for all subjects in secondary education, including standards for Dutch: the so called basic standards. In theory these basic standards will serve two aims: on the one hand they define the minimum competency level, while on the other hand they serve as an aim for teachers and students, because they define the competency level which has to be reached at the end of the third grade.

In this article the compatibility of both functions is discussed. The proposal of basic standards is discussed in the light of the consequences in terms of pass/fail decisions. Therefore the consequences of several procedures for standardsetting are related to achievements in a national assessment at the same educational level (i.e. third grade). The results seem to indicate that no matter what standards will be used, large amounts of students will fail, especially because the standards are being used in a conjunct way. The variances in achievementscores between students seem to be too large for setting standards which can function in the above mentioned manner.

'Geef mij maar werk'. Jongeren over de tijdelijke voorziening gemeentelijke werkgelegenheidsinitiatieven voor jongeren

F. MEIJERS en R. VAN WIJK
Vakgroep Andragogiek, Sectie Jongerenstudies, Rijksuniversiteit te Leiden

Samenvatting

Sedert 1 juli 1987 experimenteert een groot aantal gemeenten met de voorloper van het Jeugdwerkgarantieplan: de Tijdelijke Voorziening Gemeentelijke Werkgelegenheidsinitiatieven voor Jongeren (GWJ). Uit de hier gepresenteerde onderzoeksresultaten blijkt dat deelnemende jongeren een instrumentele houding ten aanzien van het GWJ hebben: zij verwachten door hun deelname een goede kans op een vaste baan te maken. Hun ervaringen tijdens de eerste maanden van deelname zijn voor velen van hen echter aanleiding om hieraan te gaan twijfelen, hetgeen resulteert in drop out dan wel in een expliciete vraag aan projectbegeleiders moeite te doen om hen aan een baan te helpen. Hetgeen buiten de taakstelling van de projecten valt.

Inleiding

Nog steeds zijn grote aantallen jeugdigen werkloos: in 1987 220.000 mensen onder de 25 jaar, d.i. 32% van het totaal aantal werklozen (Ministerie van Sociale Zaken en Werkgelegenheid, 1988). In zeker opzicht is de huidige situatie zorgwekkender dan de oude: de werkloosheid concentreert zich nu bij bepaalde groepen, te weten bij etnisch/culturele minderheden, laag opgeleide vrouwen en lichamelijk gehandicapten die permanent werkloos dreigen te worden (Sociaal en Cultureel Planbureau, 1989). Om ook deze groepen een kans op werk te geven heeft het kabinet het Jeugdwerkgarantieplan (JWG) ontwikkeld. Het JWG biedt alle jongeren onder de 21 jaar, wanneer ze binnen een half jaar na het verlaten van het onderwijs geen baan gevonden hebben of na beëindiging van een baan werkloos geworden zijn, een tijdelijke, boventallige (additionele) baan in de collectieve sector (rijksoverheid, provincies, gemeenten en de gepremieerde en gesubsidieerde sector). De achterliggende gedachte is dat een werkloze jongere op een garantieplaats werkervaring en het liefst ook scholing krijgt die hem of haar in staat stelt met meer kans op succes te solliciteren, vooral ook in de marktsector.

De invoering van het JWG is reeds enkele malen uitgesteld. Vanaf 1 juli 1987 hebben gemeenten echter de mogelijkheid om te experimenteren met projecten die vooruitlopen op het JWG. Daartoe is de Tijdelijke Voorziening Gemeentelijke Werkgelegenheidsinitiatieven voor Jongeren (verder aan te duiden als GWJ) in het leven geroepen. In het kader van de GWJ waren begin 1988 bijna 400 projecten goedgekeurd met in totaal 7.800 arbeidsplaatsen. Ruim 3.500 jongeren werkten op dat moment daadwerkelijk in de projecten.

In deze bijdrage presenteren we enige resultaten uit een onderzoek naar de ervaringen van aan het GWJ deelnemende jongeren. Met behulp van een gestructureerde vragenlijst zijn in het najaar van 1988 300 jongeren verspreid over 21 gemeenten geïnterviewd. De vragenlijst, een verkorte en op onderdelen aangepaste versie van de vragenlijst van Siegelaer (1988), is afgenomen door kaderleden van de Jongerenbeweging verbonden aan de FNV, die opdrachtgever van het onderzoek was. Het gevolg hiervan was dat slechts GWJ-projecten benaderd zijn in regio's waar de Jongerenbeweging een districtsbureau heeft (waarmee een landelijke spreiding overigens wel gegarandeerd was). In 16 gemeenten zijn de respondenten rechtstreeks benaderd op een in werktijd georganiseerde bijeenkomst (237 respondenten; 82% respons). In drie gemeenten kregen de GWJ-deelnemers de vragenlijst per post toegezonden en werd deze via de GWJ-projectleider doorgestuurd naar de Jongerenbeweging (57 respondenten; 35% respons). In 2 gemeenten zijn de GWJ-deelnemers individueel benaderd door kaderleden hetgeen een respons van slechts 15%

Pedagogische Studiën 1990 (67) 16-25

opleverde (6 respondenten). Ondanks deze verre van ideale procedure blijkt de onderzoeksgroep op geen enkel punt significant te verschillen van de representatieve samples uit reeds verricht onderzoek naar de GWJ van zowel Siegelaer (1988) als het Instituut voor Psychologisch Marktonderzoek (1988).

De onderzoeksgroep bestaat voor 73,9% uit vrouwen hetgeen een oververtegenwoordiging betekent in verhouding tot hun aandeel in de geregistreerde werkloosheidscijfers. Ruim tweederde (69,6%) van de door ons onderzochte jongeren is 19 of 20 jaar oud; 27,4% is jonger dan 19 en slechts 3% ouder dan 20. Iets meer dan 17% van de jongeren behoort tot een etnisch/culturele minderheid; dit is een lichte oververtegenwoordiging in verhouding tot het aandeel van de etnisch/culturele minderheden in de totale werkloosheid. Ruim de helft (55,6%) van de respondenten volgde een lagere beroepsopleiding; 18,2% volgde een MBO- en 12,8% een KMBO-opleiding. Iets meer dan de helft (53,4%) van de respondenten verliet het onderwijs met een diploma. Over het geheel genomen zijn de vrouwelijke respondenten wat hoger opgeleid: 27,8% van de meisjes (tegen 14,3% van de jongens) heeft de hogere trap van het voortgezet onderwijs (HAVO/VWO) gevolgd en 16,4% van de meisjes tegen 3,9% van de jongens heeft op dit niveau ook een diploma gehaald.

1 *Achtergronden en probleemstelling*

Er zijn de afgelopen jaren nogal wat studies verricht naar de attitudes van (werkloze) jongeren ten opzichte van – al dan niet verplichte – deelname aan diverse vormen van leer/ervaringsplaatsen (in Nederland is de meest bekende studie die van Ter Huurne, 1985). Uit al deze studies komt een merkwaardige tegenspraak naar voren. Enerzijds zeggen werkloze jongeren in meerderheid voorstander te zijn van deelname aan dergelijke projecten omdat zij deze projecten positief waarderen in termen van het opdoen van werkervaring en het verwerven van nieuwe kwalificaties. Anderzijds echter achten zij deze 'voordelen' van werk/ervaringsprojecten met name van toepassing op anderen; uit feitelijke deelnamecijfers blijkt dat minder dan 10% van de in aanmerking komende jongeren deelneemt. Daarbij blijkt dat de 'drop out' uit leer/ervaringsprojecten groot is. Veelal sluit minder dan een vijfde van de oorspronkelijke deelnemers het project op een door de projectleiding geplande en/of gewenste wijze af.

Raffe en Smith (1987) suggereren, naar aanleiding van hun onderzoek naar het Youth Training Scheme dat er sprake is van een toenemende instrumentele houding van werkloze jongeren t.a.v. leer/ervaringsprogramma's. Werkloze jongeren zouden weliswaar overtuigd zijn van het feit dat zij zonder adequate bijscholing en werkervaring weinig kans maken op het verwerven van een vaste baan (en vanuit dit perspectief een positieve attitude ten toon spreiden met betrekking tot leer/ervaringsprogramm's), maar zij zouden er tegelijkertijd aan twijfelen of deze programma's hun wel een goede kans op het verwerven van een reguliere baan geven. Er is, aldus Raffe en Smith, sprake van een dilemma dat niet logisch opgelost kan worden aangezien voor beide opties harde feiten zijn aan te dragen. Het gevolg hiervan zou zijn dat het besluit om deel te nemen c.q. deel te blijven nemen aan dergelijke programma's niet het resultaat is van een weloverwogen keuze maar afhankelijk is van een per individu verschillend samenspel van omgevingsfactoren.

In 1987 hebben we met behulp van kwalitatieve methoden onderzocht hoe 41 langdurig werkloze jongeren tussen de 20 en de 23 jaar oud (waarvan de helft deelnam aan een leer/ervaringsproject in Rotterdam en de andere helft de controlegroep vormde) omgingen met het bestaan van een dergelijk project (Meijers & Peters 1987, 1989). Het betreffende leer/ervaringsproject toonde zeer veel overeenkomsten met het GWJ. De onderzoeksresultaten deden vermoeden dat er sprake was van een dubbele instrumentalisering. Bij het besluit al dan niet deel te nemen worden jongeren geconfronteerd met het door Raffe en Smith beschreven dilemma. Degenen die besluiten om deel te nemen worden na enkele maanden opnieuw met de vraag geconfronteerd of hun deelname wel nut heeft. Het hier gepresenteerde onderzoek is opgezet om te trachten een aantal resultaten uit het Rotterdam-onderzoek kwantitatief te onderbouwen. De probleemstelling luidt als volgt: 'Hoe beoordelen deelnemers aan een GWJ-project dit project; treden er veranderingen

op in hun oordeel en zo ja, wat zijn daarvan de consequenties?'

Voordat we de onderzoeksresultaten weergeven geven we eerst een korte samenvatting van het 'Rotterdam-onderzoek'.

2 *Het Rotterdam-onderzoek: een kwalitatieve voorstudie*

Bijna alle 41 ondervraagde jongeren werkten in de eerste maanden van hun werkloosheid een tijdje via uitzendbureaus. Zij hoopten op deze manier aan vast werk te komen omdat potentiële werkgevers hen leerden kennen. Maar al snel kwamen zij er achter dat deze strategie ('wanneer je eenmaal binnen bent, dan...') voor hen absoluut niet werkte. Slechts een enkeling bleek in staat zich als uitzendkracht zodanig in de 'picture' te plaatsen dat aan haar/hem een vast dienstverband werd aangeboden. De meesten moesten werk verrichten dat inhoudelijk noch qua toekomstperspectief (te weten: een vast dienstverband) iets voorstelde. Dus stopten zij na enige tijd met werken via uitzendbureaus of beperkten zij dit tot een minimum.

De uitzichtloosheid en de daarmee samenhangende onzekerheid van hun bestaan was na 6 tot 12 maanden niet meer te negeren. Vanaf dat moment vielen de respondenten in een 'zwart gat'. Maar ook al zaten zij emotioneel en financieel zeer in de knel, toch stonden zij niet te trappelen om deel te nemen aan een leer/ervaringsplaats. Hun ervaringen op de arbeidsmarkt hadden hun namelijk geleerd dat zij, om aan vast werk te komen, tegelijkertijd moesten voldoen aan twee tegenstrijdige eisen. Aan de ene kant hadden ze ervaren dat zij te laag gekwalificeerd waren en derhalve moesten zij (her)investeren in onderwijs en scholing. Aan de andere kant wisten ze uit eigen ervaring dat het moeilijk concurreren is met schoolverlaters vanwege hun leeftijd. Opnieuw investeren in onderwijs kost tijd en dat maakte hun concurrentiepositie nog zwakker dan die al was. De meeste jongeren losten deze paradox op door te kiezen voor een meervoudige handelingsstrategie: zij investeerden in korte, arbeidsmarktrelevante cursussen en bleven tegelijkertijd zeer alert op een baan en exploiteerden daartoe alle kanalen die hun ter beschikking stonden. Het wederom deelnemen aan het reguliere onderwijs of het accepteren van een leer/ervaringsplaats werd onder meer afgewezen omdat dit een langdurige investering vraagt hetgeen betekende dat zij, zo verwachtten ze, bij het voltooien van een opleiding sowieso te oud zouden zijn om nog een kans op een vaste baan te maken.

De langdurig werkloze jongeren kozen pas voor een leer/ervaringsplaats wanneer hun ervaringen op de arbeidsmarkt hun ondubbelzinnig hadden bewezen dat zij – zelfs met de pas verworven kwalificaties via korte arbeidsmarktrelevante cursussen – er nooit in zouden slagen op eigen kracht een vaste baan te vinden. Zij kozen pas voor deelname wanneer alle 'bronnen', waaruit hoop kon worden geput op het verkrijgen van een vaste baan, opgedroogd waren. Daarmee was deze keuze in de ogen van vooral de mannelijke respondenten een strategie van de (bewezen) onmacht. Zij stonden echter vrijwel allemaal positief ten opzichte van dergelijke projecten, vooral als 'laatste kans' voor hen 'die het echt niet meer weten'.

De deelnemers bleken het project vooral te waarderen als een 'veredeld uitzendbureau'. Dat wil zeggen: als een instelling die hen in staat stelde met meer succes de strategie van 'als je eenmaal binnen bent, dan...' te volgen dan commerciële uitzendbureaus dat naar hun gevoel deden. Zij kregen, naar zij hoopten en zoals hun door het project in de wervingsfolders ook werd beloofd, de kans om zich gedurende ten minste een jaar op een zinvolle arbeidsplaats waar te maken in de ogen van een potentiële werkgever. Zij ontwikkelden een strategie die het best getypeerd kan worden als een strategie van de offerbereidheid. Zij hoopten dat zowel hun omgeving als potentiële werkgevers het feit dat zij bereid waren tegen een geringe vergoeding (te weten 5 gulden per dag) 'normale' arbeidsdagen te maken positief waardeerden.

Na enkele maanden werken op een leer/ervaringsplaats bekroop hun echter het vermoeden dat hun hoop op een vaste arbeidsplaats wel eens misplaatst zou kunnen zijn. Zij merkten dat zij door hun collega's op het werk niet voor 'vol' werden aangezien en dat werken op een leer/ervaringsplaats stigmatiserende effecten had. Potentiële werkgevers beschouwden hen namelijk, wanneer zij solliciteerden, als 'kneusjes'. Daarbij kwam nog dat zij op de leer/ervaringsplaats veelal

behandeld werden als een 'gewone' uitzendkracht of zelfs slechter. Het gevolg hiervan was dat de tevredenheid met het project (i.c. met de scholing, de beloning en de inhoud van het werk) sterk terugliep en een aantal deelnemers ophield. De blijvers gingen veel meer begeleiding vragen. Zij verwachtten met name dat de projectbegeleiders voor hun aan actieve arbeidsbemiddeling gingen doen. Dit was echter – onder de bestaande regelgeving – onmogelijk zodat ook de blijvers uiteindelijk de waarde van hun investering ernstig ter discussie gingen stellen.

3 *De onderzoeksresultaten*

3.1 *Instroom en aanvangsmotivatie*

Het merendeel (60,4%) van de deelnemers aan de GWJ-projecten komt het project binnen via het Gewestelijk Arbeidsbureau of de Gemeentelijke Sociale Dienst. Slechts 11,1% komt via familieleden en/of vrienden en 13,1% via advertenties en/of wervingsfolders in aanraking met een GWJ-project. Deelname aan een GWJ-project is echter geheel vrijwillig. Op dit punt verschilt het GWJ van het JWG dat deelname verplicht stelt. Kijken we naar Tabel 1, dan valt op dat er sprake is van ongelijke kansen op deelname aan een GWJ-project. De hoger gekwalificeerde mannen zijn het minst lang werkloos, voordat zij een GWJ-project instromen. Zij profiteren in feite het snelst van deze nieuwe maatregel. Ruim 60% van deze groep is naar eigen zeggen minder dan een half jaar werkloos geweest. Zij zijn binnen enkele maanden nadat zij van school kwamen via GSD/GAB het project binnengeloodst. Het aantal hoog gekwalificeerde mannen is echter te gering om verregaande conclusies te kunnen trekken. Hoger gekwalificeerde vrouwen zijn in meerderheid eerst een half tot een heel jaar werkloos voor zij instromen.

Tabel 1 *Duur van de werkloosheid voor deelname aan project in relatie tot onderwijsniveau en geslacht in % (N=289)*

Duur werkloosheid	Mannen		Vrouwen		Totaal	
	Lager	Hoger	Lager	Hoger	Lager	Hoger
	niveau onderwijs*					
<1/2 jaar	29.2	63.6	22.9	33.3	24.8	38.0
1/2-1 jaar	43.1	27.3	49.0	60.0	47.2	54.9
>1 jaar	27.7	9.1	28.1	6.7	28.0	7.1
(Absoluut	65	11	153	60	218	71)
	p ≦0,05 (mannen) en ≦0.01 (vrouwen)					

* Onder lager niveau verstaan wij de lagere trap voortgezet onderwijs inclusief het KMBO; onder hoger niveau de hogere trap voortgezet onderwijs.

Bijna de helft van de lager gekwalificeerde mannen en vrouwen (d.w.z. zij die slechts de lagere trap van het voortgezet onderwijs of het KMBO gevolgd hebben), voor wie de maatregel toch in eerste instantie bedoeld is omdat zij de meeste kans lopen (langdurig) werkloos te worden, zijn tussen 6 en 12 maanden werkloos geweest terwijl 27,7% van de lager gekwalificeerde mannen en 28,1% van de lager gekwalificeerde vrouwen zelfs langer dan een jaar zonder werk was. Het lijkt erop dat de keuze om deel te nemen aan een GWJ-project slechts voor de lager opgeleide jongeren een 'strategie van de onmacht' is. Zij voelen zich waarschijnlijk, gelet op de cultuur van de lagere sociale milieus waartoe zij vrijwel allen behoren, op 19 of 20 jarige leeftijd op de drempel van volwassenheid staan en voelen zich gefrustreerd wanneer zij deze drempel niet kunnen overschrijden omdat zij geen werk hebben. Dat zij later instromen zou dan verklaard kunnen worden door het feit dat zij langer dan andere leeftijdsgenoten naar werk zoeken. Daarmee hebben ze echter in meerdere opzichten een achterstand opgelopen vergeleken met hun leeftijdsgenoten.

Mannen profiteren sneller dan vrouwen en hoger gekwalificeerden weer sneller dan lager gekwalificeerden van de GWJ; hetzelfde kan gezegd worden van autochtonen vergeleken met allochtonen. Vergeleken met hun aandeel in de totale werkloosheid zijn allochtone jongeren niet alleen oververtegenwoordigd bij de instroom maar ook bij het vroegtijdig vertrek uit de projecten. Ruim de helft (55%) van de autochtone deelnemers neemt langer dan 6 maanden aan een project deel tegen nog geen 40% van de allochtone deelnemers. Het is, gelet op de werkloosheidscijfers onder allochtone groepen, onwaarschijnlijk dat meer allochtone dan autochtone deelnemers het project verlaten hebben omdat zij een baan gekregen hebben.

Het is opmerkelijk dat juist de jongeren met de slechtste arbeidsmarktperspectieven minder lang deelnemen aan een project. Het

lijkt er op dat de GWJ, evenals dat trouwens met andere arbeidsmarktinstrumenten het geval is (zie hiervoor bijv. Dijkhuis-Potgieser, 1984), met name ten goede komt aan jongeren met wat betere arbeidsmarktperspectieven.

Uit Tabel 2 blijkt dat jongeren in overgrote meerderheid aan een GWJ-project deelnemen op basis van dezelfde argumenten die politici gebruik(t)en wanneer zij het bestaan van deze projecten rechtvaardigen. Het opdoen van werkervaring (41,7%) of van scholing en werkervaring (14,9%) in combinatie met het geloof dat men via deelname gemakkelijker aan een vaste baan kan komen (33,2%): precies dit is de boodschap die de projecten naar werkloze jongeren uitstralen. En deze boodschap pikken jongeren op.

Tabel 2 *Redenen van deelname in % (N = 295)*

Het opdoen van werkervaring	41.7
Om gemakkelijker aan een vaste baan te komen	33.2
Het krijgen van scholing in combinatie met werkervaring	14.9
Uit verveling	6.8
Ik zat thuis opgesloten en ontsnapte via het project	2.4
Ik werd onder druk gezet door GAB en/of GSD	1.0

Deze 'aanvangsmotivatie' kan echter van enige kanttekeningen voorzien worden. Dit wordt duidelijk wanneer we kijken naar de antwoorden op de vraag: 'Denk je na afloop van het project kans te hebben op een baan?'. Bijna één derde (31,5%) beantwoordt deze vraag met een 'ja', een vrijwel even groot aantal (29,2%) met een 'nee' en 39,3% ten slotte met een 'ik weet het niet'. Blijkbaar gelooft minder dan één derde de 'boodschap' van het project (en van de politici) echt. Daaruit kan men afleiden dat de meerderheid op hoop van zegen investeert in de GWJ: 'baat het niet, het schaadt toch zeker ook niet' lijkt het parool. Dit wordt ook duidelijk uit de gegevens van Tabel 3.

Slechts één kwart van de deelnemers is uit zichzelf de mening toegedaan dat het beter zou zijn wanneer zij zich eerst nog verder kwalificeerden. Drie kwart (76,8%) wil een gewone betaalde baan.

Ook bij de GWJ-deelnemers lijkt sprake van een dubbele strategie om de paradox 'te laag gekwalificeerd – te oud' op te lossen. De gegevens uit Tabel 4 laten zien dat ruim één derde van de deelnemers van meet af aan blijft solliciteren. Het aantal solliciterenden loopt verder omhoog tot 63% van degenen, die langer dan 6 maanden aan de GWJ deelnemen. Leeftijd en opleidingsniveau beïnvloeden dit beeld niet.

Tabel 3 *Voorkeur op moment van interview in % (N = 298)*

ik wil een gewone betaalde baan hebben	76.8
ik wil eerst meer werkervaring opdoen	11.1
ik wil een normale vakopleiding volgen	8.1
ik wil naar een dagschool	3.7
ik wil werkloos blijven	0.3

Tabel 4 *Sollicitatiegedrag gerelateerd aan de duur van de deelname in % (N = 291)*

	minder dan 3 maanden	3-6 maanden	langer dan 6 maanden
solliciteert wel	37.7	56.6	63.0
solliciteert niet	62.3	43.4	36.1
(Absoluut	53	83	155)
		$p \leq 0.01$	

Van alle jongeren die blijven solliciteren naar een baan doet 80% – onafhankelijk van de duur van deelname – dit omdat het werk in de GWJ slechts van tijdelijke aard is; slechts 6,7% solliciteert omdat men het met de GWJ-baan verdiende loon te gering vindt. Dit is opmerkelijk omdat de deelnemende jongeren, en dan met name de jongens, niet tevreden zijn met het loon dat zij verdienen. We komen hierop direct nog terug. Wat uit de gegevens duidelijk wordt is dat jongeren geen tijdelijk werk willen. Zij willen een 'normale' baan oftewel: vast werk waarop je een toekomst kunt bouwen (zie ook Meijers, 1989).

3.2 *De aard van het GWJ-werk*

In het GWJ blijken deelnemers, onafhankelijk van hun kwalificatie-niveau, vooral administratief werk te verrichten: 37,7% van de mannen en 56,2% van de vrouwen doen dit werk. De gegevens van Tabel 5 laten daarbij een lichte tendens zien tot een geslachtsspecifieke verdeling van het aangeboden werk. Ruim één vijfde van de mannelijke deelne-

mers (allen laag gekwalificeerd) verricht onderhouds- of ander technisch werk tegen 1% van de vrouwelijke deelnemers. Deze laatsten (en dan met name laag gekwalificeerden onder hen) verrichten behalve administratief werk vooral schoonmaak- en huishoudelijk werk, werk in keuken of kantine of verzorgend werk.

Tabel 5 *Verrichte werkzaamheden in relatie tot geslacht in % (N = 296)*

	Mannen	Vrouwen	Totaal
administratief werk	37.7	56.2	51.4
technisch werk	5.2	0.5	1.7
onderhoudswerk	20.8	0.5	5.7
werk op een bouwplaats	1.3	0.0	0.3
keuken/kantine/ schoonmaakwerk/ huishoudelijk werk/ verzorging	16.9	29.7	26.3
postzaken/ archiefwerk	2.6	3.7	3.4
bezorgen/ vervoerswerk	1.3	0.9	1.0
anders	14.3	8.7	10.1
(Absoluut	77	219	296)

Deze verdeling van werkzaamheden geeft de mannelijke deelnemers een iets betere kans op doorstroming naar een reguliere baan. De groei van de werkgelegenheid vindt nl. de laatste tijd met name plaats in de technische beroepen (in de bouw en metaalnijverheid). De omvang van het administratieve en het huishoudelijke en verzorgende werk daarentegen is de laatste jaren afgenomen. In deze situatie zal, naar verwachting, de komende jaren weinig verandering komen (OSA, 1988).

Het werk dat op de garantiebanen verricht wordt, is naar de mening van 41,4% van de deelnemende jongeren niet additioneel. Het is, met andere woorden, werk dat – in de definitie van De Vries en Ter Huurne (1987, p. 1) – wettelijk verplicht is en/of in het jaar voorafgaande aan de inschakeling van werklozen uitbesteed werd en/of verricht werd door vast personeel terwijl er in hetzelfde voorafgaande jaar bezuinigd werd op dit werk en/of dat, nadat het niet meer verricht wordt door werklozen, (wederom) door regulier personeel verricht wordt.

Hoe hoger de deelnemer gekwalificeerd is, des temeer is zij of hij van mening dat het werk dat verricht wordt niet additioneel is: 66,7% van de jongeren met een HAVO/VWO-opleiding zonder diploma en 79,5% HAVO/-VWO-ers met diploma meent dat het werk dat zij verrichten anders ook zou gebeuren. Deze jongeren, 72 in getal (24% van het sample), verrichtten overigens allen administratief werk. Maar ook de helft van de 228 laag gekwalificeerde deelnemers is van mening dat het werk dat zij nu verrichten vroeger ook gebeurde.

Dat de deelnemers waarschijnlijk een realistisch beeld hebben met betrekking tot het al dan niet additioneel zijn van het door hen verrichte werk maken De Vries en Ter Huurne (1987) duidelijk. Zij concluderen op basis van hun onderzoeksgegevens dat 80% van het werk dat werklozen binnen gemeentelijke afdelingen en diensten verrichten, niet additioneel is en dus ten koste gaat van normale banen. Zij spreken de verwachting uit dat jongeren in het kader van het JWG terecht zullen komen op arbeidsplaatsen, waarop bezuinigd is en waarvoor velen van hen juist een opleiding volgen. Daarbij worden verdere bezuinigingen mogelijk: "Het werk wordt immers toch gedaan en daardoor is er geen noodzaak tot afbouw van taken" (o.c., p. 80).

Crommelin (1985) constateerde in zijn onderzoek naar het functioneren van leer/ervaringsplaatsen in het bedrijfsleven, dat het zeer moeilijk is additioneel werk aan te bieden dat tegelijkertijd wel voldoende mogelijkheden biedt op het verwerven van relevante werkervaring: "Of men voert stipt de regels van het plan uit, in welk geval de deelnemer vaak onvoldoende nuttige ervaring opdoet, of men doet een oogje dicht en neemt als uitgangspunt dat de deelnemer in ieder geval (...) zinvol bezig moet zijn, waarbij er naar wordt gestreefd dat het zo veel mogelijk additioneel werk zal betreffen". Het lijkt erop – als we afgaan op de mening van de deelnemers op dit punt – dat het creëren van zinvol additioneel werk in de gemeentelijke organisatie eveneens moeilijk is en dat eenzelfde 'oplossing' in praktijk wordt gebracht als door Crommelin beschreven.

Gemeten aan het belang dat politici eraan hechten lijkt er weinig ruimte in de GWJ-projecten te zijn voor om-, her- of bijscholing. Bijna 40% van de deelnemers zegt dat zij door het project niet gestimuleerd zijn om

Tabel 6 *Verwachte hulp van het project bij het vinden van een baan en hierbij gekregen hulp, onderscheiden naar duur van de deelname in % (N = 293)*

	Hulp verwacht			Hulp krijgend		
	<3 mnd	3-6 mnd	>6 mnd	<3 mnd	3-6 mnd	>6 mnd
Ja	51.9	64.3	73.1	33.3	21.2	35.7
Nee	20.4	12.9	9.0	16.7	49.4	37.0
Weet niet	27.8	22.9	17.9	50.0	29.4	27.3
(Absoluut	54	85	154	54	85	154)
		$p \leq 0.05$			$p \leq 0.01$	

zich om-, her- of bij te scholen. Hoe lager het kwalificatieniveau, des te minder men – naar eigen zeggen – gestimuleerd wordt. Van de deelnemers die LBO of MAVO gevolgd hebben zegt 57,4% gestimuleerd te worden verdere scholing te volgen tegen 78,4% van de deelnemers die HAVO, VWO of MBO gevolgd hebben. Naarmate men langer aan een project deelneemt, wordt men slechts een beetje meer gestimuleerd om verdere scholing te volgen. Eén derde van de deelnemers, die langer dan 6 maanden deelnemen, zegt niet gestimuleerd te worden tegen 39,8% van de totale groep. Vrijwel de helft (152 respondenten) ontvangt op dit moment, naar eigen zeggen, geen enkele vorm van scholing.

3.3 *De invloed van de duur van deelname op de houding ten aanzien van het GWJ-project*

In het 'Rotterdam-onderzoek' constateerden wij dat deelnemers aan een leer/ervaringsproject na enkele maanden gingen twijfelen aan hun hoop om, via een 'strategie van de offerbereidheid', aan een vaste baan te komen. Zij merkten dat zij noch door collega's op hun werk noch door potentiële werkgevers voor 'vol' werden aangezien. Daarbij kwam dat zij op hun leer/ervaringsplaats veelal behandeld werden als een 'gewone' uitzendkracht of zelfs slechter. Het gevolg daarvan was dat de tevredenheid met de scholing, de beloning en de inhoud van het werk sterk terugliep en een aantal deelnemers ophield. De blijvers gingen meer begeleiding vragen (vooral een actieve arbeidsbemiddeling van de kant van de projectleiding).

Tabel 6 laat zien dat er verschillen optreden in de hulp die men verwacht en die men naar eigen zeggen krijgt tussen degenen die korter dan drie maanden, drie tot zes maanden en langer dan zes maanden deelnemen aan een GWJ-project. Overigens verschillen deze groepen niet wat betreft kwalificatieniveau en leeftijd. Van de deelnemers die korter dan drie maanden deelnemen verwacht 51,9% hulp bij het vinden van een baan. Dit percentage loopt op tot 73,1% bij degenen die langer dan zes maanden deelnemen. In beide gevallen ligt de hulp, die de deelnemers naar eigen zeggen in dit opzicht krijgen, ver onder hun verwachting.

Tabel 6 laat verder zien dat er in de periode tussen de derde en de zesde maand een daling optreedt van het aantal deelnemers dat zegt hulp te krijgen terwijl er tegelijkertijd een zeer scherpe stijging optreedt van het aantal deelnemers dat zegt geen hulp te krijgen. Daarbij groeit de hulpvraag (i.e. de verwachting dat men hulp zal krijgen) gestaag. Deze gegevens zijn in overeenstemming met het beeld dat uit het Rotterdam-onderzoek naar voren komt.

Van dit beeld afwijkend is het gegeven dat 35,7% van degenen die langer dan zes maanden deelnemen zeggen hulp te krijgen tegen 21,2% van degene die tussen de drie en zes maanden deelnemen. Blijkbaar zien de projectbegeleiders, meer dan in Rotterdam het geval was, voor zich de mogelijkheid de deelnemers te helpen bij het vinden van een baan.

Er zijn nog andere aanwijzingen dat er tussen de derde en de zesde maand van deelname sprake is van toenemende twijfel aan de zinvolheid van deelname en van toenemende onzekerheid over de eigen kansen op werk. Allereerst het punt van de beloning. Op de vraag, of men rond kan komen van het verdiende loon, wordt verschillend geantwoord. Een meerderheid van de jongens (56%) is van mening dat zij van dit loon niet kunnen rondkomen; daarentegen vindt 54,1% van de meisjes de beloning voldoende: zij kunnen er van

rond komen. Differentiëren we echter naar de duur van de deelname, dan verandert er nogal wat zoals Tabel 7 laat zien.

Tabel 7 *Oordeel over het verdiende loon, gerelateerd aan de duur van de deelname in % (N = 291)*

	<3 mnd	3-6 mnd	>6 mnd
Ik kan er van rondkomen	66.7	44.6	46.1
Ik kan er niet van rondkomen	27.8	49.4	51.9
Weet niet	5.6	6.0	2.0
(Absoluut	54	83	154)
		p ≤0.01	

In het Rotterdam-onderzoek bleek dat, zolang deelnemers het gevoel hadden dat zij een goede investering pleegden door deel te nemen aan het project, de hoogte van de beloning voor hen niet zo belangrijk was. De beloning mocht zelfs (te) laag zijn: dat paste in hun strategie van 'offerbereidheid'. Na enige maanden gingen velen echter twijfelen aan de zinvolheid van de investering met als een van de gevolgen dat de tevredenheid met de beloning sterk afnam.

De gegevens van Tabel 7 lijken er op te wijzen dat in de GWJ-projecten eenzelfde proces plaats vindt. Tussen de derde en de zesde maand loopt het percentage deelnemers dat zegt te kunnen rondkomen van het verdiende loon, terug van 66,7 naar 44,6% terwijl bijna de helft (en na zes maanden zelfs meer dan de helft) van de jongeren aangeeft niet rond te kunnen komen.

Tabel 8 *Mening over scholing gerelateerd aan de duur deelname in % (N = 152)*

	<3 mnd	3-6 mnd	>6 mnd
(Zeer) goed	78.6	53.8	59.6
Matig/slecht	21.4	46.2	40.4
(Absoluut	14	39	99)
		p ≤0.01	

Een volgende aanwijzing dat er sprake is van toenemende twijfel en onzekerheid, vinden we in de teruglopende tevredenheid met de aangeboden scholing. De gegevens uit Tabel 8 maken duidelijk, dat ook op dit punt sprake is van een scherpe daling in tevredenheid met de aangeboden scholing tussen de derde en de zesde maand. Overigens zegt slechts de helft van de ondervraagde deelnemers enige vorm van scholing te krijgen.

Een vierde aanwijzing vinden we wanneer we kijken naar de kansen die deelnemers zich geven om na afloop van het project een vaste baan te vinden. Tabel 9 maakt duidelijk dat één derde aan het project begint in de verwachting na afloop van het project kans te maken op een vaste baan; meer dan de helft twijfelt echter reeds op dat moment. Terwijl het aantal deelnemers, dat zich kansen toedicht, het gehele project door gelijk blijft, neemt vooral het aantal deelnemers dat zich geen kansen op een vaste baan geeft, tussen de derde en de zesde maand fors toe (van 15,1% tot 36,9%).

Tabel 9 *Inschatting van de kans op vast werk na afloop van het project, in relatie tot de duur van de deelname in % (N = 291)*

	<3 mnd	3-6 mnd	>6 mnd
Ik maak kans op vast werk	30.2	31.0	32.5
Ik maak geen kans op vast werk	15.1	36.9	29.9
Weet niet	54.7	32.1	37.7
(Absoluut	53	84	154)
		p ≤0.05	

Hiervoor hebben we laten zien (Tabel 4) dat deelnemers een dubbele strategie hanteren: zij nemen deel aan het project doch blijven solliciteren. Bijna 80% van degenen die solliciteren geeft, onafhankelijk van de duur van hun deelname aan de GWJ, als voornaamste reden om te solliciteren het feit dat het werk dat ze nu in het kader van het GWJ verrichten slechts tijdelijk werk is. Voor slechts 9,3% van de deelnemers, die al langer dan zes maanden deelnemen, is de hoogte van het loon een reden om weer te gaan solliciteren; voor hen die korter deelnemen is dit ueberhaupt geen reden. Dit is opmerkelijk omdat na enkele maanden deelname minder dan de helft van de deelnemers zegt rond te kunnen komen van het verdiende loon (Tabel 7).

Niet alle deelnemers kiezen evenwel voor een dubbele strategie. Eén derde van alle deelnemers zet alle kaarten op het project en besluit niet te solliciteren. Het gaat hier vooral

om jongeren met de minste kansen op de arbeidsmarkt; om deelnemers dus die laag gekwalificeerd zijn. De LBO/MAVO-ers zonder diploma solliciteren het minste (46,2% solliciteert); de HAVO/VWO/MBO-ers met diploma het meeste (74,4%). De gegevens uit Tabel 10 geven een goede indicatie van wat er bij deze groep aan de hand is.

Tabel 10 *Redenen om niet te solliciteren in % (N = 109)*

Ik denk dat ik geen kans op een baan maak	12.8
Ik hoop via het project aan vast werk te komen	64.2
Ik wil (nog) geen vast werk	18.3
Er is toch geen werk	4.6

Ruim 60% van alle deelnemers die niet solliciteren, hoopt via het project aan vast werk te komen. Dit betekent dat ook de niet-solliciteerders die al langer dan zes maanden aan het project deelnemen, in meerderheid niet solliciteren omdat zij slechts via het project een kans op vast werk denken te maken.

4 *Conclusie*

Bovenstaande resultaten lijken Raffe en Smiths instrumentaliseringsthese te ondersteunen. De onderzoeksopzet maakte het niet mogelijk na te gaan of een instrumentele houding ten aanzien van leer/ervaringsprojecten reeds bestaat voordat men deelneemt aan dergelijke projecten. Duidelijk is wel dat jongeren deelnemen aan een project omdat ze er een instrumenteel nut van verwachten (te weten een betere kans op een vaste baan). Zij willen deze vaste baan liefst direct; zij nemen alleen deel aan een GWJ-project omdat ze niet aan een vaste baan kunnen komen.

Tijdens de deelname verandert de houding naar het project aantoonbaar. Het oordeel over de beloning en over de aangeboden scholing wordt slechter terwijl men tegelijkertijd meer hulp gaat vragen bij het zoeken en vinden van een baan. We hebben er vanaf gezien om de jongeren te vragen of zij twijfelen aan het nut van hun deelname. In het vooronderzoek dat we in Rotterdam uitvoerden bleek dat twijfel aan de eigen 'investering' bij een groot aantal deelnemers omsloeg in een fanatiek onderschrijven van de projectdoelstellingen. Dergelijke psychologische processen zijn met behulp van een gestructureerde vragenlijst niet op te sporen.

Nergens zijn evenwel indicaties te vinden dat het veranderende oordeel over verschillende belangrijke aspecten van het project veroorzaakt wordt door andere factoren dan twijfel aan het nut van deelname. Nader onderzoek, waarbij ook jongeren moeten worden betrokken die voortijdig uit het project verdwijnen (gegevens hierover ontbreken op dit moment nog volledig), zal op dit punt echter definitief uitsluitsel moeten verschaffen.

Geraadpleegde literatuur

Crommelin, R. B., Ervaringen met Jeugdplan bij ESSO Nederland. *De Gids*, 1985, nr. 3, 58-70

Dijkhuis-Potgieser, H. I. E., *Evaluatie experimentele arbeidsprojecten voor jeugdige werklozen*. 's-Gravenhage: Staatsuitgeverij, 1984

Huurne, A. G. ter, *Arbeidsprojecten voor werkloze jongeren. Wat zeggen zij er zelf van?*. Tilburg: IVA, 1985

Huurne, A. G. ter & A. de Vries, *Uitkeren of belonen?* Een onderzoek naar de motieven voor en de effecten van inschakeling van werklozen binnen gemeentelijke afdelingen en diensten. Tilburg: IVA, 1987

Instituut voor Psychologisch Marktonderzoek, *Uitvoering van de Tijdelijke Voorziening Gemeentelijke Werkgelegenheidsinitiatieven voor Jongeren*. Een interimrapport. Rotterdam: IPM, 1988

Kooistra, S. & H. Kroft (red.), *Maatregelen voor werklozen; wat werkt en wat niet?* Evaluatieonderzoek uit zeven Westeuropese landen bijeengebracht. Rijswijk: Ministerie van WVC, 1986

Meijers, F. & E. Peters, *'t Schiet niet op. . . . Werkloze jongeren over leer/ervaringsplaatsen*. Eindverslag van een onderzoek in opdracht van de Gemeente Rotterdam, het G.A.B./Rotterdam en het Ministerie van WVC. Leiden: Rijksuniversiteit, 1987

Meijers, F. & E. Peters, Leer/ervaringsplaatsen: een alternatief voor werkloosheid? In: J. van Damme & J. Dronkers (red.), *Jongeren in school en bedrijf*. Lisse: Swets & Zeitlinger, 1989

Meijers, F. (red.), *Jongeren op weg. Eerste onderzoeksverslag uit het project 'Jongeren en Arbeid'*. Leiden: Rijksuniversiteit, 1989

Ministerie van Sociale Zaken en Werkgelegenheid, *Rapportage Arbeidsmarkt 1987*. 's-Gravenhage: SDU, 1988

Organisatie voor Strategisch Arbeidsmarktonderzoek, *OSA-Rapport 1988: Arbeidsmarktperspectieven*. 's-Gravenhage: SDU, 1988

Raffe, D. & P. Smith, Young people's attitudes to YTS: the first two years. *British Educational Research Journal*, 1987, nr. 3, 241-260
Siegelaer, E., *Jeugdwerkgarantieplan: garantie of tijdelijke voorziening?* Utrecht: CNV, 1988
Sociaal en Cultureel Planbureau, *Sociaal Culturele Verkenningen 1990*. 's-Gravenhage: SDU, 1989

Curricula vitae

F. Meijers (1950) studeerde westerse sociologie te Nijmegen en was achtereenvolgens werkzaam in Nijmegen (Sociologisch Instituut) en Amsterdam (Subfaculteit Andragologie). Sinds 1986 werkzaam bij de Vakgroep Andragogiek, Sectie Jongerenstudies van de Rijksuniversiteit te Leiden. Speciale belangstelling voor de problematiek rond de overgang van school naar werk. Schreef onder andere 'Van ambachtsschool tot LTS' (1983) en 'De toekomst van de volwasseneneducatie' (1987).

R. van Wijk (1961) studeerde in 1987 af binnen de Vakgroep Andragogiek, Sectie Jongerenstudies van de Rijksuniversiteit te Leiden. Voerde vervolgens – samen met Meijers – twee contractonderzoeken uit en werkt momenteel als beleidsondersteunend ambtenaar bij de Gemeente Amersfoort, alwaar hij zich speciaal bezig houdt met werkgelegenheidsprojecten voor langdurig werkloze jongeren.

Adres: Vakgroep Andragogiek, Sectie Jongerenstudies, Rijksuniversiteit te Leiden, Postbus 9555, 2300 RB Leiden

Manuscript aanvaard 21-8-'89

Summary

Meijers, F. & R. van Wijk. 'I'd rather like a job.' *Pedagogische Studiën*, 1990, *67*, 16-25

Since July 1987 many municipalities experiment with the official forerunner of a forthcoming nation-wide Youth Training Scheme: the Temporary Municipal Employment Initiatives for Youngsters (MEI). In this article it is shown that youngsters have an instrumental attitude towards MEI: they expect that joining MEI gives them a good chance to get a regular job. Their experiences during the first months of participation, however, give rise to serious doubts about this chance. This, in turn, leads to drop out or to an explicit demand that the project supervisors will do more for them with respect to job-finding. Under the present regulations this is impossible.

Het effect van articulatie op het onthouden van geschreven woorden bij dove kinderen*

M. W. SCHAPER**
Effatha, Christelijk Instituut voor Doven, Voorburg

Samenvatting

In dit artikel wordt ingegaan op de vraag hoe dove kinderen het best kunnen worden geïnstrueerd in hun leeswoordenschatontwikkeling. Verslag wordt gedaan van een trainingsexperiment waarin het effect is nagegaan van twee oefencondities. Een eerste groep leerlingen leerde geschreven woorden te associëren met betekenis zonder specifieke leerinstructie. Een tweede groep leerlingen werd geïnstrueerd om tijdens oefening hardop te lezen. Uit gegevens van dit onderzoek blijkt dat onder de conditie 'vrij-leren' dove kinderen in het onthouden van geschreven woorden gebruik maken van een overwegend visuele strategie. Met de oefenconditie 'hardop-lezen' wordt bewerkstelligd, dat in het onthouden van woordbeelden gebruik wordt gemaakt van articulatorische specificaties van woorden. De twee oefencondities laten evenwel geen onderscheiden effect zien in vaardigheden die de ontwikkeling van een orthografisch lexicon betreffen. Wel levert de oefenconditie 'hardop-lezen' een meer effectief memoriseren van woordreeksen op, hetgeen mogelijk van belang is voor het begrijpend lezen van tekst.

* Het hier vermelde onderzoek maakt deel uit van een onderzoeksproject, dat dankzij financiële ondersteuning van het Nederlands Comité voor Kinderpostzegels tot stand kwam.

** De auteur is dank verschuldigd aan de leerkrachten en leerlingen die in het onderzoek hebben geparticipeerd, aan P. Reitsma voor zijn methodologische adviezen en de gegevensverwerking en M. Wissink voor haar aandeel daarin, aan A. Dees, I. de Graaff en G. de Kleer voor hun medewerking aan de uitvoering van het onderzoek.

1 *Inleiding*

Prelinguaal dove[1] kinderen verwerven doorgaans een zeer beperkte leeswoordenschat (Hatcher & Robbins, 1978; IJseldijk, 1989; Kyle, 1980; Walter, 1978) en worden onder meer door dit lexicaal tekort ernstig belemmerd in hun leesontwikkeling. Kenmerkend verschil tussen het leren lezen van dove en horende kinderen is, dat bij doven aan het leesleerproces geen spreektaalverwerving vooraf gaat. Dove kinderen ontwikkelen hun mondelinge taalvaardigheden uitsluitend of dominant langs visuele weg en daartoe is de grafische weergave van de taal door haar eenduidig en permanent karakter (vgl. hiermee de modaliteiten spraakafzien en vingerspelling) het belangrijkste middel. In 'oraal' georiënteerd onderwijs aan doven, d.w.z. onderwijs dat zich richt op het ontwikkelen van mondelinge en schriftelijke communicatievaardigheden met een nadrukkelijk uitsluiten van manuele communicatievormen, is het schrift aanvankelijk ook het belangrijkste medium om inzicht en kennis omtrent de omringende wereld over te dragen. Recentelijk zijn in Nederland de meeste scholen overgegaan op een 'totaal-communicatief' georiënteerd onderwijs, waarbij naast mondelinge en schriftelijke communicatie ook het gebaren-maken en vingerspellen als een natuurlijk deel van de interactie met het dove kind worden geaccepteerd en gestimuleerd. Dit levert onder meer voor de leesontwikkeling het grote voordeel, dat semantisch kan worden aangesloten op een reeds aanwezig gebarenlexicon. Ondanks voornoemd verschil hebben beide communicatieve scholingsconcepten gemeen, dat het leren lezen simultaan verloopt met de (spreek)taalverwerving. Een belangrijke implicatie is, dat doven in het leren lezen minder gebruik kunnen maken van de alfabetische aard van het schriftsysteem, i.c. een strategie van verklanken naar een reeds aanwezig spreektaalbezit niet kunnen benutten. Met betrekking tot het verwerven van een leeswoordenschat heeft dit tot gevolg, dat dove kinderen elk afzonderlijk woord als aparte vorm

moeten onthouden en deze woorden rechtstreeks leren associëren met betekenis.

De gangbare aanpak van voornoemd leesleerproces dat bij dove kinderen reeds op 3 à 4 jarige leeftijd wordt ingezet, is vooral gericht op het zich eigen maken van de juiste betekenisinformatie van woorden. Aanvankelijk wordt uitgegaan van hoog-imaginaire begrippen (bv. 'pop', 'bed', 'fiets'), waarbij associaties worden gelegd met concrete voorwerpen en/of plaatjesafbeeldingen. Problemen rijzen echter in het begrijpend leren lezen van laag-imaginaire begrippen als bv. 'blij', 'vakantie' en 'plagen'. Hier is methodisch gebruikelijk om aansluiting te zoeken bij beleving van leerlingen, waarbij vocale en/of manuele uitingen door de leerkracht worden omgezet in juiste bewoording en vervolgens in tekst vastgelegd. Deze schriftelijke neerslag vormt enerzijds de leerstof voor de (spreek)taalontwikkeling. Daarnaast levert het de leesstof voor het leesonderwijs, waarbij de feitelijke leesleerhandeling aanvankelijk gericht is op een langs globaal-intuïtieve wijze weer herkennen van de grafische informatie op grond van reeds aanwezige conceptualisaties (zie voor een meer gedetailleerde beschrijving o.a. Ewoldt & Hammermeister, 1986: de 'language-experienceapproach' en Van Uden, 1982, 1988: het 'ideo-visueel lezen'). Deze holistische benadering van het leren lezen staat dove leerlingen echter toe een zeer globaal beeld te vormen omtrent woorden. Vooral bij toenemende woordenschat brengt deze globale orthografische kennis het risico, dat verwarrende effecten kunnen ontstaan in woordherkenning en daarmee, dat woordbetekenis associaties moeilijker worden onthouden.

Bij normaalhorende leerlingen wordt vaardigheid in vlot en nauwkeurig woordherkennen als een belangrijke voorwaarde gezien voor het leren lezen (Perfetti & Lesgold, 1977, 1979; Stanovich, 1982). Tevens blijkt dat kennis van de grafemische structuur van woorden van groot belang is voor het verwerven van die vaardigheid in woordherkenning (Barron, 1986; Reitsma, 1983, 1985). Dit heeft aanzet gegeven om te beproeven of dove kinderen in het onthouden van woord-betekenis associaties gebaat zijn met aanvullende oefeningen ter bevordering van een meer accurate orthografische kennis van woorden. Bij de opzet van het beoogde behandelingsprogram is gebruik gemaakt van een tweetal uitgangspunten. Ten eerste wordt er van uitgegaan dat voor het vergroten van woordspecifieke kennis oefening met losse woorden het meest effectief is. Deze opvatting vindt empirische ondersteuning in onderzoek van Ehri en Roberts (1979) en Ehri en Wilce (1980), waarin naar voren is gebracht dat door het lezen van woorden in zinnen de syntactische en semantische informatie van woorden beter wordt aangeleerd en met het lezen van losse woorden de orthografische structuur beter wordt onthouden. Een tweede uitgangspunt van behandeling is, dat opslag en associatie van lexicale informatie in het lange termijn geheugen bevorderd worden door een inprentingsmethodiek (zie Van der Leij, 1985). Naar praktische uitwerking impliceert dit een herhaald aanbieden van leesleerinhoud en het geven van directe feedback op het leerproces. De wijze, waarop dove kinderen het best kunnen worden geïnstrueerd in het onthouden van woorden vormt onderwerp van studie en wordt via een aantal experimentele oefencondities in de onderwijspraktijk geëvalueerd. In deze bijdrage wordt verslag gedaan van een trainingsexperiment, waarin het effect is nagegaan van articuleren als middel om het onthouden van woorden te bevorderen.

2 *Mogelijk belang van articuleren in het onthouden van geschreven woorden*

Hoewel veel dove kinderen via grafeemfoneem omzetting[2] de spreekvorm van geschreven woorden leren reconstrueren, komt voor hen zoals reeds gezegd met een articulatorische aanpak in het leren lezen van (nieuwe) woorden geen betekenisinformatie vrij. Desalniettemin lijkt voor dove lezers het toepassen van spraakcodering van belang in het opnemen en verwerken van geschreven informatie. Zo is bij dove adolescenten gebleken, dat met articulatorisch coderen tijdens het memoriseren van schriftelijk aangeboden woordreeksen meer informatie kan worden vastgehouden dan met een puur visueel of manueel coderen en dat de mate waarin van spraakcodering gebruik wordt gemaakt positief samenhangt met leesontwikkeling (Conrad, 1979; Hanson, 1982; Lichtenstein, 1983). De verklaring van deze samenhang is gezocht in het feit, dat het voor begrijpend lezen nood-

zakelijk is om woorden voor korte tijd vast te houden in het werkgeheugen om interpretatie van betekenisrelaties tussen woorden in een zin(sdeel) toe te laten, een proces dat bij normaalhorende lezers wordt bevorderd door het fonologisch coderen van geschreven informatie (zie o.a. Baddeley, Eldridge & Lewis, 1981; Kleiman, 1975; Slowiczak & Clifton, 1980).

Het articulatorisch coderen van woorden kan mogelijk ook bijdragen aan een betere vorming van orthografische kennis van woorden. Deze veronderstelling wordt ondersteund door bevindingen uit eerder onderzoek, waarin werd nagegaan hoe dove leerlingen in de leeftijd van 7 tot 13 jaar geschreven woorden onthouden in associatie met betekenis (Schaper & Reitsma, 1986). Geconcludeerd werd, dat jonge leerlingen praktisch altijd gebruik maken van een visuele strategie. Vanaf 9-jarige leeftijd ontwikkelen sommige kinderen echter een voorkeur voor een articulatorisch strategiegebruik, hetgeen in vergelijking tot een visuele aanpak allengs een meer effectieve strategie blijkt: de associaties woordbeeld-betekenis worden sneller en beter geleerd. Als verklaring is aangevoerd, dat de dove lezer door het articuleren wordt gedwongen tot een betere analyse van de specifieke lettercombinaties van woorden en daardoor de grafemische structuur van woorden gemakkelijker onthoudt (zie Bus, 1984 voor een vergelijkbaar aangetoond effect van het hardop-lezen bij normaalhorende kinderen). Het voordeel van een betere analyse van woorden is hier echter niet exclusief toe te schrijven aan een articulatorisch coderen. Eenzelfde effect wordt geopteerd voor het gebruik van vingerspellen, waarbij geschreven lettercombinaties in een één-op-één relatie worden omgezet in een manueel-alfabetische codering (Van Dijk, 1985; Hirsch-Pasek & Treiman, 1982; Hirsch-Pasek, 1987). Daarnaast zou ook met (over)schrijven van woorden de aandacht kunnen worden gevestigd op woordstructurele kenmerken.

Indien een nauwkeurige kennis van de grafemische structuur een gunstig effect heeft op het onthouden van woorden, dan kan dit wellicht via onderwijs worden bevorderd door dove leerlingen tijdens leesoefeningen systematisch gebruik te leren maken van articulatie, vingerspellen of schrijven. Van de geboden alternatieven vergt het verklanken van woorden van jonge dove kinderen relatief veel inspanning. Desalniettemin wordt aan articulatie als strategie in het onthouden van woorden een groot gewicht toegekend. Ten eerste wordt in het Nederlandse onderwijs aan doven, ook waar vorm is gegeven aan standpunten van 'totale communicatie', een hoge prioriteit gesteld aan de ontwikkeling van mondelinge communicatie-vaardigheden. Articulatorisch coderen tijdens het leren lezen brengt mogelijk het voordeel, dat samen met grafemische kenmerken ook uitspraakspecificaties van woorden worden aangeleerd. Een tweede weliswaar speculatieve overweging in het accentueren van articulatie in het onthouden van losse woorden is het mogelijk voordeel van een transfer effect op codegebruik in het onthouden van woordreeksen. Zoals reeds is toegelicht blijkt spraakcodering hier het meest effectief en van invloed op het begrijpend lezen van tekst. Om voornoemde redenen wordt aan vingerspellen of schrijven als mogelijk middel om het onthouden van woorden te bevorderen voorlopig een complementaire functie toegeschreven, waar het articulatorisch coderen in gebreke blijft.

3 *Opzet en uitvoering van het onderzoek*

3.1 *Vraagstelling en opzet*

Doel van het onderzoek is na te gaan of jonge dove kinderen zijn te beïnvloeden in het gebruik van spraak in het onthouden van geschreven woorden en wat daarvan het effect is op leesontwikkeling. Daartoe leerden dove kinderen gedurende vier maanden wekelijks een setje woorden met betekenisinformatie associëren, met de instructie om tijdens oefenen hardop te lezen. Om eventuele effecten van articulatie te kunnen bepalen is een vergelijk gemaakt met een controle-conditie, waarin leerlingen zonder specifieke leerinstructie de woordsetjes leerden associëren met betekenis. Deze conditie van 'vrij-leren' biedt een benadering van de wijze, waarop dove kinderen in de gangbare onderwijspraktijk – weliswaar via het herhaald lezen van woorden in zin of tekst – woorden leren onthouden. De verwachting is dan ook dat in de controleconditie gebruik wordt gemaakt van een overwegend visuele strategie, analoog aan reeds eerder gerapporteerde bevindingen (Schaper & Reitsma, 1986).

Om vast te kunnen stellen of met hardop

lezen wordt bewerkstelligd dat articulatorische informatie (de motorisch-kinesthetische spraakindrukken) tijdens het leren wordt benut en in 'vrij-leren' vooral gesteund wordt op grafisch-visuele informatie van woorden is de leerlingen na afloop van de training de reeds ontwikkelde en beproefde taak uit voornoemd onderzoek van Schaper en Reitsma voorgelegd. In deze proef worden leerlingen gevraagd om setjes (pseudo)woorden te associëren met betekenisinformatie. Een onderscheid is aangebracht in enerzijds woordsetjes, waarbij de woorden onderling sterk gelijkende uitspraak hebben, maar visueel duidelijk verschillen (conditie a). Daarnaast zijn setjes samengesteld uit woorden die visueel erg op elkaar lijken maar qua uitspraak verschillen (conditie v). De rationale van deze taak is dat bij visueel strategiegebruik in het onthouden van woorden vooral problemen zullen ontstaan in conditie v ten gevolge van visuele verwarring. Bij een articulatorisch strategiegebruik zal door gelijkenis in het motorisch-kinesthetisch patroon het onthouden van woorden in conditie a relatief moeilijk zijn.

Betreffende de vraag naar eventuele bijzondere leereffecten van spraakcodering is om te beginnen bepaald of woord-betekenis associaties sneller en beter worden onthouden. Daartoe is de benodigde oefenfrequentie bezien voor het bereiken van een vastgelegd criterium voor beheersing, als ook de retentie van het geleerde een week na oefening. Verder is met behulp van een tweetal taken nagegaan of er een betere orthografische kennis van woorden tot stand wordt gebracht. Ten eerste is gebruik gemaakt van een woordherkenningstaak, waarbij een plaatje moet worden gematcht met één uit vier visueel sterk gelijkende schrijfwijzen van een doelwoord. Indien het hardop lezen een positief effect heeft op orthografische kennis van woorden, dan zou dit tot uiting kunnen komen in een accurate en vlotte uitvoer van de proef. De tweede taak betrof een 'dictee' van aangeleerde woorden. Omdat dove kinderen bij woordspelling doorgaans gebruik maken van een visueel-orthografische strategie (Wimisner & Arnold, 1986), zal een eventueel positief effect van hardop lezen op woordstructurele kennis hier tot uiting komen in een betere spelling van aangeleerde woorden. Ten slotte is onderzocht of met articuleren tijdens oefening met losse woorden wordt bewerkstelligd, dat spraakcodering meer wordt toegepast in het memoriseren van woordreeksen. Op grond van hetgeen in de vorige paragraaf staat vermeld kan dit tot uiting komen in een betere geheugencapaciteit.

In de opzet van het onderzoek is gebruik gemaakt van het 'terugtrekkingsdesign'. Na een basislijn periode van twee weken (controle-instructie) is een eerste experimentele fase gevolgd van zes weken (controle resp. experimentele instructie), waarna in de daaropvolgende twee weken alle leerlingen volgens controle-conditie oefenden om na te gaan in hoeverre sprake was van generalisatie of terugval van de experimentele groep op het basisniveau. Daarop volgde een tweede experimentele fase van zes weken, weer afgesloten met twee weken oefenen onder de controleconditie.

3.2 *Onderzoeksgroep*

In het onderzoek werden alle 20 kinderen betrokken, die bij aanvang van het onderzoek in het eerste en tweede leerjaar van het SO van Effatha zaten (6 resp. 14 leerlingen). De leerlingen zijn prelinguaal doof zonder complex meervoudige handicaps, met een leeftijd variërend tussen 6 en 8 jaar. Allen hebben het thans in Nederland gebruikelijke 'totaalcommunicatief' georiënteerde kleuteronderwijs voor doven gevolgd en via spreekonderwijs goeddeels geleerd om schriftbeelden 'technisch' om te zetten in spraak.

Voorafgaand aan het leerexperiment zijn de leerlingen op grond van een aantal voor het leren lezen relevant geachte variabelen zoveel mogelijk gelijk verdeeld over de experimentele en controlegroep. Beide groepen komen gemiddeld goed overeen naar leerjaar, leeftijd, intelligentie, mate van gehoorverlies, articulatievaardigheid en capaciteit van nonverbaal simultaan/successief geheugen (statistische toetsing voor verschillen in gemiddelden leverde voor alle variabelen een $F<1$).

3.3 *Materialen*

Oefenmateriaal werd een week voor iedere training van een woordset samengesteld. De keuze van woorden geschiedde op grond van een wekelijks vooronderzoek naar de bekendheid van 15 à 25 geschreven woorden, door de leerlingen aan te geven met gebaar en/of plaatjesassociatie. Hieruit werden die woor-

den geselecteerd, die nagenoeg voor alle kinderen onbekend waren. Voor woordkeuze is geput uit een reeks samenhangende verhalen om met het oog op betekenisgeving de woorden in een context te kunnen aanbieden. Dit uitgangspunt beperkte echter het hanteren van criteria die de moeilijkheidsgraad van het aanleren van woorden kunnen beïnvloeden (bijv. mate waarin woorden onderling visueel of articulatorisch verschillen, bekendheid met en imaginairiteit van betekenis). Alle woorden waren samengesteld uit vier tot acht letters. Woordsoort betrof vooral het zelfstandig naamwoord, werkwoord (infinitief) en bijvoeglijk naamwoord; verdeling van woordsoort over de sets fluctueerde.

Betekenisinformatie van geselecteerde woorden werd door een illustrator in tekening gebracht. Naast een zgn. vertelplaat waarin de te leren begrippen binnen verhaalcontext aan de orde konden worden gesteld, zijn ook 'losse' voorstellingen van woorden gemaakt. In deze afbeeldingen werd waar mogelijk een concrete representatie van het begrip gegeven; bij lage imaginairiteit betrof be'tekende' de gebarencode, gepresenteerd door een aan de vertelling ontleend personage.

In de eerste helft van het leerexperiment (tot de elfde week) kregen leerlingen uit eerste en tweede leerjaar per week respectievelijk 4 en 6 woorden aangeboden, in de tweede helft van het onderzoek werd dit aantal uitgebreid tot 6 resp. 8 woorden. Tijdens het oefenen werden de woorden met een aanbiedingstijd van 5 seconden gepresenteerd op de monitor van een computer. Betekenisassociatie werd gelegd door bijhorende plaatjesvoorstellingen aan te laten wijzen op een kaart met alle afbeeldingen van het woordsetje, aangevuld met twee aan dezelfde context verbonden plaatjes (1e en 2e leerjaar resp. 6 en 8 keuzealternatieven). Van elke kaart zijn 4 versies gemaakt met afwisselende volgorde van plaatjes. Omdat in de eerste helft van het onderzoek bleek dat niet relevante plaatjes zelden worden aangewezen is in de tweede helft van het onderzoek alleen gebruik gemaakt van de bij de woordset behorende plaatjes.

3.4 *Procedure*

Het wekelijks leren van een woordset startte met een klassikale introductie van de woorden via een verhaal en bijhorende vertelplaat. De vertelling werd in gebaar en woord gepresenteerd, waarbij de te leren woordbetekenissen nadrukkelijk aan de orde werden gesteld.

Het oefenen van de woorden vond individueel plaats. Elke trainingsweek omvatte 3 oefenzittingen op achtereenvolgende dagen. Per zitting werd met maximaal 8 leertrials (ca. 15 minuten) de set woorden aangeleerd. Wanneer de leerling op de eerste dag het maximale aantal leertrials nodig bleek te hebben werd een tweede dag geoefend. In de andere gevallen werd alleen gedurende de eerste en derde dag gewerkt.

De eerste oefenzitting begon met een introductie, waarin de aan het verhaal ontleende begrippen werden gerelateerd aan de losse plaatjesafbeeldingen en deze vervolgens aan de geschreven woorden. Kinderen die leerden onder de conditie 'hardop lezen' werden gevraagd de woorden te verklanken; zo nodig werd hier gecorrigeerd op grafeem-foneem omzetting en/of op zorgvuldig spreken. Leerlingen in de controle-conditie werden vrij gelaten in de manier waarop ze de associaties legden. In de hierop volgende leertrials werden de leerlingen geïnstrueerd één voor één de op de monitor aangeboden woorden in associatie te brengen met de bijbehorende plaatjes. Kinderen in de experimentele conditie werden ook hier geïnstrueerd hardop te lezen en tijdens het oefenen zo nodig gewezen op fouten in grafeem-foneem omzetting en/of gestimuleerd tot zorgvuldig spreken. Aan kinderen in de controle-conditie werden deze eisen niet gesteld. Aanbieding van de volgorde van de woordreeks varieerde per leertrial. Daarnaast werd na steeds twee woord-aanbiedingen ook een kaart met gewijzigde plaatjesvolgorde gebruikt. Na ieder woord-plaatje associatie werd door de onderzoeksassistent feedback gegeven en zo nodig op fouten gecorrigeerd. Oefening werd afgebroken bij twee achtereenvolgende correcte trials voor alle woorden. Per woord-plaatje associatie werd de goed/fout score geregistreerd. Tevens werden observaties met betrekking tot code-gebruik bij oefenen (articulatie, vingerspelling, gebaren maken) genoteerd. Uit de scoringlijsten zijn als samenvattende cijfers opgemaakt hoeveel trials (gesommeerd over 2 of 3 oefenzittingen) een leerling nodig had in de training en het totaal aantal woorden dat tijdens de oefenzittingen fout werd aangewezen.

Eén week na afsluiting van een oefenweek werd in een retentietoets nagegaan hoeveel

van de in de training geleerde woorden nog her/gekend werden. De procedure was conform oefenen echter zonder op fouten te corrigeren. In deze natoetsen werden twee trials van alle woorden aangeboden. De score van de retentietoets betrof het aantal woorden dat in beide trials correct werd geassocieerd.

Binnen het onderzoek vond op gezette momenten tussen- en natoetsing plaats. Na afloop van de eerste en tweede experimentele fase (na resp. 8e en 16e week) is de leerlingen de woordherkenningstaak voorgelegd. Het woordmateriaal voor de herkenningstaken werd gevormd door woorden die in de retentietoetsen van respectievelijk de eerste en tweede reeks experimentele oefenperiode door minstens 80% van de leerlingen goed was onthouden (proportie niet onthouden woorden werd over de experimentele en controle-groep gelijk gehouden). Het aantal items voor 1e en 2e leerjaar was voor de eerste afname respectievelijk 12 en 18 en voor de tweede afname 18 en 24. Het aantal goede antwoorden en de benodigde tijd werden geregistreerd.

In de 6 weken van de tweede experimentele oefenperiode is direct na afloop van de laatste oefenzitting het dictee van aangeleerde woorden afgenomen. Instructie was om bij het tonen van de plaatjes het bijhorend woord op te schrijven. Het aantal volledig correct gespelde woorden werd gescoord.

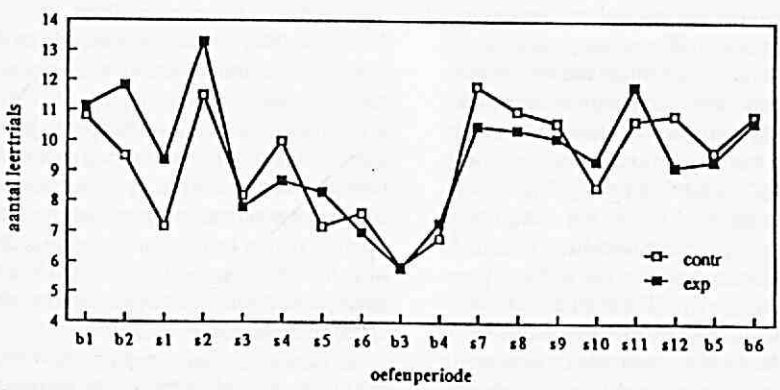
Voor aanvang van het trainingsexperiment en na afloop van de eerste en tweede experimentele fase vond afname plaats van de proef 'geheugen voor woordseries'. Leerlingen werd gevraagd in reeks aangeboden woorden te onthouden, om vervolgens in dezelfde volgorde de bijbehorende plaatjes aan te geven uit 4 afbeeldingen. Er is gebruik gemaakt van reeds in kleuteronderwijs aangeleerde, zeer goed bekende woorden. De proef omvatte in totaal 6 series van respectievelijk 2, 3 en 4 woorden. Voor de aanbieding werd gebruik gemaakt van een computer (aanbiedingstijd per woord 2 seconden met een interval van 1 seconde tussen de woorden). De proef werd afgebroken bij 4 achtereenvolgende fouten. De score is het aantal correct weergegeven series.

Ten slotte vond na afloop van de laatste basislijnperiode de proef 'strategiegebruik in het onthouden van woorden' plaats. Leerlingen werd gevraagd 4 setjes van 3 pseudowoorden te associëren met afbeeldingen van dieren: twee setjes samengesteld onder conditie a (articulatorisch overeenkomend, visueel verschillend) en twee onder conditie v (visueel overeenkomend, articulatorisch verschillend). De woordsetjes zijn in alternerende volgorde van a en v conditie aangeboden. Afname was conform het trainingsonderzoek, echter zonder speciale leerinstructie en met gewijzigde procedure van feedback (bij fouten werd na afloop van een leertrial correctie geboden). Oefenen werd beëindigd na 2 opeenvolgende foutloze trials en gemaximaliseerd tot 8 oefentrials. Het totaal aantal benodigde trials voor a-sets resp. v-sets werd gescoord.

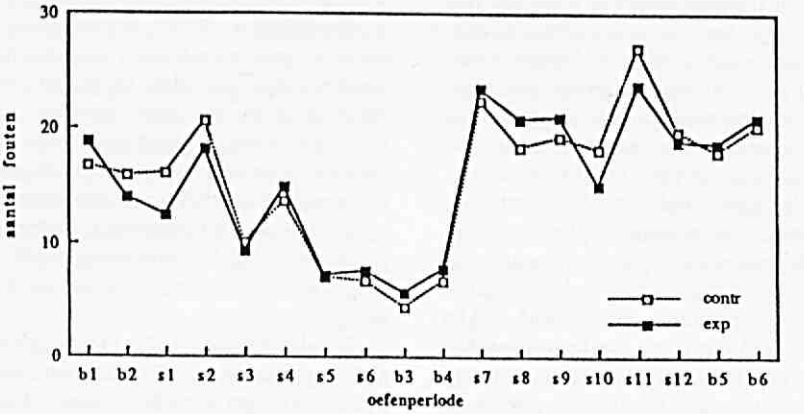
4 *Resultaten*

4.1 *Resultaat van oefenproces*

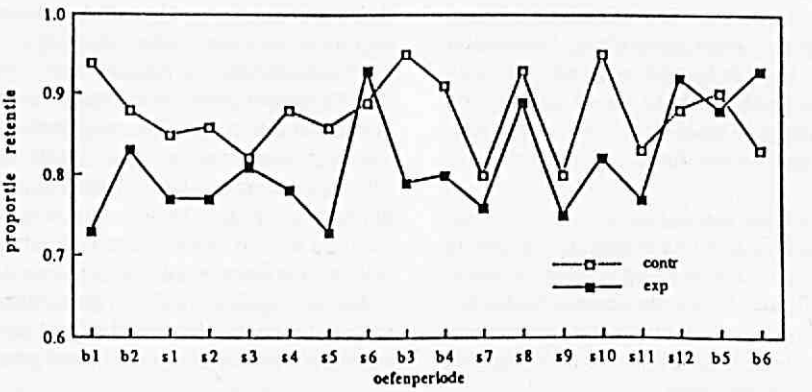
Per groep en per week zijn de gemiddelden van het totaal aantal leertrials, gesommeerd



Figuur 1 Aantal leertrials per oefenweek



Figuur 2 *Aantal fouten tijdens oefenen*



Figuur 3 *Proportie retentie per oefenperiode*

over de 2 of 3 oefenzittingen per trainingsweek, weergegeven in Figuur 1. De in totaal 18 trainingsweken zijn verdeeld in 3×2 basislijnperioden (b1 - b6) en 2×6 experimentele trainingsweken (s1 - s12). Uit deze figuur is duidelijk op te maken dat het aantal benodigde leertrials een nogal grillig verloop toont; zoals verwacht is de moeilijkheidsgraad van de te leren set woorden niet voor iedere oefenweek gelijk. De vrij abrupte toename van aantal leertrials aan het begin van de tweede experimentele fase (s7) is zeer waarschijnlijk veroorzaakt door de uitbreiding van het aantal te leren woorden op dat moment. Voorts is duidelijk in Figuur 1 te zien dat de twee groepen leerlingen in vrijwel dezelfde mate worden beïnvloed door de relatieve moeilijkheidsgraad van de te leren woorden. Belangrijk voor de onderzoeksvraagstelling is de constatering dat de experimentele groep niet significant afwijkt van de controlegroep; statistische toetsingen wijzen uit dat er géén verschillen zijn tussen beide groepen ($F's < 1$). Wel is het aantal leertrials sterk afhankelijk van de oefenweek en dus van de specifieke woorden die dan geleerd moeten worden ($F(17,306) = 5.39$, $p < .001$). Tussen de basislijn en de experimentele trainingsperioden is geen significant verschil geconstateerd.

Het totaal aantal fouten tijdens elke oefenweek is eveneens per groep gemiddeld en weergegeven in Figuur 2. Duidelijk is te zien dat het aantal fouten tijdens het oefenen voor beide groepen leerlingen niet erg uiteenloopt; ook hier wijst statistische toetsing uit dat er geen verschillen zijn tussen de experimentele en controlegroep ($F's < 1$).

In Figuur 3 is per groep en per week het gemiddelde weergegeven van de proportie goed herkende woorden in de een week na de laatste

Tabel 1 *gesommeerde scores 'articulatie' per oefenweek en groep*

	b1	b2	s1	s2	s3	s4	s5	s6	b3	b4	s7	s8	s9	s10	s11	s12	b5	b6
contr.	0	.3	.3	2.4	2.4	4.2	4.3	3.9	3.3	4.3	0.9	1.3	1.9	2.2	2.6	3.3	2.0	2.6
exp.	0	0	10	10	10	10	10	10	10	9.3	10	10	10	10	10	10	10	10

trainingszitting afgenomen retentietest. Uit de grafiek komt de tendens naar voren dat de controlegroep meer woorden onthoudt dan de experimentele groep. Dit verschil is echter reeds aanwezig voordat de experimentele manipulatie zijn beslag heeft gekregen. Statistische analyse maakt duidelijk dat er over het geheel genomen geen enkel verschil is tussen de experimentele en controle leerlingen in de proportie herkende woorden ($F<1$). Een uitzondering vormt de eerste basislijnweek, waar de controlegroep significant meer woorden weet te onthouden ($F(1,62) = 4.93$, $p = .03$).

Per leerling zijn tijdens iedere oefenzitting observaties met betrekking tot codegebruik genoteerd. Gedurende de experimentele fase is vanzelfsprekend de experimentele groep volgens instructie volop aan het articuleren. Maar de vraag kan worden gesteld, welk gedrag de controlegroep vertoont en hoe de experimentele groep zich gedraagt tijdens de 2e en 3e basislijnweek waar de leerlingen niet gevraagd werden te articuleren. Terwijl in de eerste basislijnperiode niet of nauwelijks werd gearticuleerd is voor de experimentele groep daarna het articuleren een niet meer weg te cijferen gedrag – ook in de tweede en derde basislijnperioden. Met betrekking tot de controlegroep valt op, dat enkele leerlingen spontaan in enigermate hardop gaan lezen, vooral gedurende de eerste leertrials van een oefenzitting. Om een vergelijk mogelijk te maken met mate van articuleren in de experimentele groep, werd het articuleren geklassificeerd in 'af en toe' (0.3 punt), 'veelvuldig' (0.7 punt) en 'consequent' (1.0 punt), een scoring die slechts een globale totaalindruk weergeeft. Per oefenweek en groep is het articulatorisch codegebruik gesommeerd en weergegeven in Tabel 1.

Het gebruik van vingerspelling is in beide groepen verwaarloosbaar. Slechts bij één leerling (controlegroep) is enkele keren het vingerspellen van de beginletter van een aangeboden woord geconstateerd. Wel wordt in beide groepen door veel leerlingen na verloop van oefening het geschreven woord gekoppeld aan de gebarencode, alvorens de bijhorende plaatjesvoorstelling aan te wijzen.

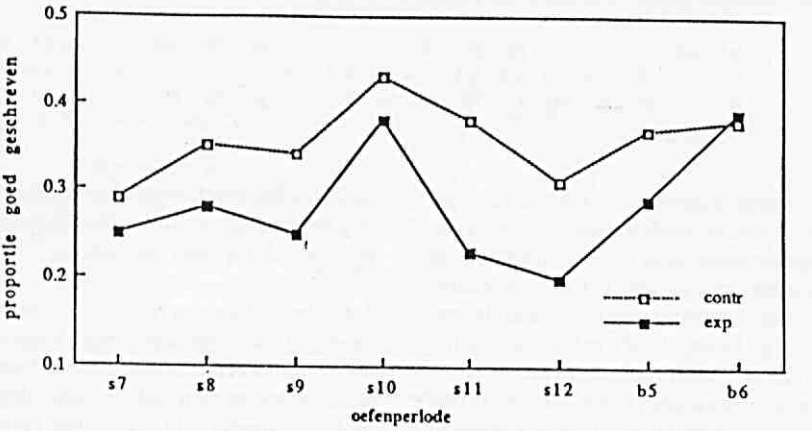
4.4 *Resultaten van tussen- en natoetsen*

De proef 'woordherkenning', afgenomen na eerste en tweede experimentele fase laat bij eerste afname zien dat de controlegroep gemiddeld minder tijd voor deze taak behoeft dan de experimentele groep (128 versus 168 s; $F(1,18) = 4.43$, $p = .05$). Het aantal goed gekozen woorden in deze test is echter niet verschillend voor beide groepen (12.5 versus 10.5; $F(1,18) = 1.3$, n.s.). Analyse van de resultaten van de tweede afname van de woordherkenningstoets geeft geen enkel significant verschil tussen de groepen (tijd: $F(1,18) = 1.07$, n.s.; aantal goed: $F<1$).

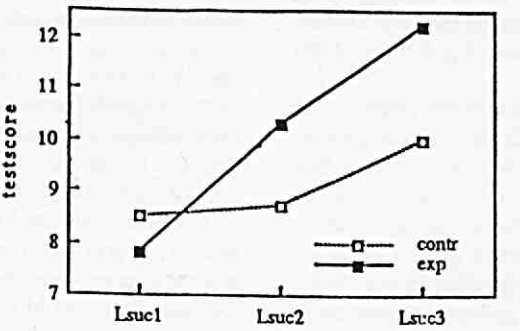
Resultaat van het 'dictee' is uitgedrukt in proporties goed geschreven woorden en in Figuur 4 weergegeven per groep en oefenweek. Globaal genomen blijkt slechts ongeveer een derde van de geleerde woorden correct te worden geschreven. De controlegroep levert iets betere spellingprestaties ($M = .357$) dan de experimentele groep ($M = .283$), maar van statistische significantie is geen sprake ($F<1$).

De proef 'geheugen voor volgorde van woorden' is in totaal driemaal afgenomen. In Figuur 5 is per testafname de gemiddelde score voor elke groep weergegeven. Uit de grafiek valt op te maken dat de experimentele groep beduidend en systematisch meer vooruitgaat in prestaties dan de controlegroep. Een variantie-analyse van deze gegevens met twee factoren (de tussenpersoons conditiefactor en de binnenpersoons herhaalde meting) laat zien dat er geen algemeen effect van oefenconditie is ($F<1$), een behoorlijk systematisch effect van het meetmoment c.q. de test ($F(2,36) = 21.43$, $p<.001$) en een interactie tussen de experimentele conditie en de herhaalde meting ($F(2,36) = 6.75$, $p<.01$).

Van de proef 'strategiegebruik in het onthouden van woorden' is het aantal oefentrials benodigd om respectievelijk een articulatorisch (S-a) en een visueel overeenkomend (S-v)



Figuur 4 *Proportie goed geschreven woorden direct na training*



Figuur 5 *Scores voor de 3 afnames 'geheugen voor woordreeksen'*

setje woorden te leren geanalyseerd op systematische effecten van soort woordset (a versus v) en oefenconditie (controle- versus experimentele groep). De variantieanalyse gaf aan dat beide hoofdeffecten niet significant ($F's<1$) waren, maar de interactie tussen woordset en oefenconditie wel ($F(1,18)=7.62$, $p=.01$). Uit de gemiddelden per groep (S-a/S-v waren 5.0/7.8 en 7.6/4.4 resp. voor controle- en experimentele groep) blijkt dat terwijl de controlegroep beduidend meer moeite heeft de visueel gelijkende woordsets aan te leren, de experimentele groep daarentegen moeite heeft met het leren van de articulatorisch gelijkende woordsets.

5 *Conclusie en discussie*

De lijn van de vraagstelling volgend leert het huidig onderzoek ten eerste, dat jonge dove kinderen zijn te beïnvloeden in het gebruik van spraakcodering tijdens het leren lezen van woorden. De uitkomst van de proef 'strategiegebruik in het onthouden van woorden' laat duidelijk zien, dat leerlingen die geïnstrueerd zijn om tijdens oefenen hardop te lezen gebruik maken van een overwegend articulatorische aanpak. Daarentegen wordt bij leerlingen die tijdens de training hun eigen leerstrategie konden bepalen een overwegend visuele aanpak aangetroffen, overeenkomstig de bevinding van eerder onderzoek (Schaper & Reitsma, 1986). Ook de paar leerlingen die onder de conditie 'vrij-leren' spontaan in enigermate overgingen op hardop lezen blijken in het onderzoek naar strategie-gebruik een dominant visuele aanpak toe te passen in het onthouden van woorden. Voor dit onderscheiden effect van 'geïnstrueerd' en 'spontaan' hardop lezen kan een tweetal verklaringen worden aangevoerd. Ten eerste kan een verschil in oefenfrequentie van invloed zijn geweest. Leerlingen die geïnstrueerd zijn om

hardop te lezen doen dit consequent gedurende de gehele oefenperiode; onder de conditie 'vrij-leren' wordt door de betreffende leerlingen gedurende een beperkter deel van training hardop gelezen. Een tweede mogelijke verklaring is dat leerlingen die geïnstrueerd zijn hardop te lezen tevens – veelvuldig – correctieve feedback kregen op de kwaliteit van verklanken. Leerlingen onder de controleconditie ontvingen een dergelijke feedback niet. Dit kan tot gevolg hebben dat eerstgenoemde geschreven woorden beter en wellicht meer bewust relateren aan spraakcodering.

Aannemende dat onder de twee oefencondities een overwegend articulatorische respectievelijk visuele strategie in het onthouden van woorden wordt ontwikkeld of toegepast, is in dit onderzoek geen evidentie gevonden voor een onderscheiden effect van dergelijke strategieën op leesleerontwikkeling. Tegen de oorspronkelijke verwachting in worden woordbeeld-betekenis associaties niet sneller geleerd, noch in retentiemeting beter onthouden indien geoefend wordt onder gebruikmaking van een articulatorische strategie. De rationale van deze veronderstelling was, dat de lezer door het articuleren wordt gedwongen tot een betere analyse van de specifieke lettercombinaties van te leren woorden en daardoor de grafemische structuur van woorden gemakkelijker onthoudt. De uitkomsten van de proeven waarin de orthografische kennis van aangeleerde woorden is nagegaan (de woordherkenningstaak en het dictee) laten evenwel zien dat ook hier geen differentieel effect optreedt tussen beide oefencondities. Opgemerkt zij dat voor het trainingsonderzoek woordsetjes zijn gebruikt die tamelijk willekeurig – geput uit verhaalcontext – zijn samengesteld, daarmee een zo goed mogelijke afspiegeling gevend van de gangbare onderwijspraktijk. Dientengevolge worden deze woordsets niet of nauwelijks gekenmerkt door specifiek articulatorische resp. visuele verwarringsmoeilijkheden. Hoewel de moeilijkheidsgraad van de te leren set woorden nogal kan verschillen, wordt onder voornoemde conditie van contextgebonden woordkeuze geen differentieel effect aangetroffen tussen een articulatorische en visuele strategie in het leren lezen.

Het ontbreken van enig bijzonder effect van spraakcodering op het verwerven van een orthografisch lexicon lijken de bevindingen van het reeds eerder aangehaalde onderzoek van Schaper & Reitsma (1986) enigszins te weerspreken. Dat onderzoek gaf aanleiding te veronderstellen dat het toepassen van articulatie in vergelijking tot een visuele benadering een meer effectieve strategie is in het onthouden van woorden. Het onderscheiden effect van strategiegebruik werd echter alleen aangetroffen bij leerlingen uit het 6e en 7e leerjaar en niet bij jongere kinderen. Het ontbreken van effect van een articulatorisch strategiegebruik op leesleerontwikkeling in het huidige onderzoek zou wellicht kunnen worden opgevat als leeftijd- of leesleerfase-specifiek. Een verklaring zou namelijk kunnen zijn, dat vorming van spraakcodes voor jonge dove kinderen verhoudingsgewijs meer aandacht en inspanning vergt, waardoor bij toepassing van een dergelijke strategie de algehele lettersamenstelling van woorden enigszins uit het oog wordt verloren.

Het enig onderscheiden effect van de verschillende oefencondities betreft de uitkomsten van de proef 'geheugen voor woordseries'. In vergelijking tot de controlegroep laat de experimentele groep in de drie afnames een beduidende en systematische vooruitgang zien in het op reeks onthouden van woordseries. Een mogelijke interpretatie van deze ontwikkelingstendens is dat in de experimentele groep relatief meer gebruik wordt gemaakt van een articulatorisch coderen dan in de controlegroep. Deze veronderstelling is afgeleid van de in paragraaf 2 besproken onderzoeksresultaten bij dove adolescenten (Conrad, 1979; Hanson, 1982; Lichtenstein, 1983), waaruit blijkt dat spraakcodering in vergelijking tot een visueel of manueel codegebruik een meer effectieve strategie is in het onthouden van woordsequenties. De aanwijzing dat in de experimentele en controlegroep een respectievelijk articulatorische en visuele strategie in het onthouden van losse woorden wordt ontwikkeld c.q. toegepast laat een dergelijke interpretatie zonder meer toe. Het mogelijk belang van dit positief effect van articulatie op de opnamecapaciteit voor geschreven woorden in het korte-termijn-geheugen is zoals reeds werd toegelicht, dat daarmee het begrijpend lezen van tekst wellicht gunstig wordt beïnvloed.

Noten

1. Doofheid die bij geboorte of door ziekte aanwezig is, voor de spontane taalontwikkeling normaliter op gang komt.
2. In plaats van het bij normaalhorenden veel gebruikte 'fonologisch (de)coderen' wordt bij dove kinderen vaak gesproken van een 'articulatorisch (de)coderen', daarmee benadrukkend dat niet de acoustische maar de motorisch-kinesthetische eigenschappen van spraakcodering de belangrijkste bron voor informatieverwerking is.

Literatuur

Baddeley, A., M. Eldridge & V. Lewis, The role of subvocalization in reading. *Quarterly Journal of Experimental Psychology*, 1981, *33*, 439-545.

Barron, R.W., Word recognition in early reading: A review of the direct and indirect access hypotheses. *Cognition*, 1986, *24*, 93-119.

Bus, A.G., *Leesproblemen en instructiemethoden*. Lisse: Swets & Zeitlinger, 1984.

Conrad, R., *The deaf schoolchild: Language and cognitive function*. London: Harper & Row, 1979.

Dijk, J.P.M. van, *Kritische beschouwingen over het leren vingerspellen van dove kinderen*. Paper gepresenteerd op een lezing te Effatha, Zoetermeer, februari 1985.

Ehri, L.C. & K.T. Roberts, Do beginners learn printed words better in context or in isolation? *Child Development*, 1979, *50*, 675-685.

Ehri, L.C. & L.S. Wilce, Do beginners learn to read function words better in sentences or in lists? *Reading Research Quarterly*, 1980, *4*, 451-476.

Ewoldt, C. & F. Hammermeister, The language-experience approach to facilitating reading and writing for hearing impaired students. *American Annals of the Deaf*, 1986, *131*, 271-275.

Hanson, V.L., Short term recall by deaf signers of American Sign Language: implications of encoding strategy for order recall. *Journal of Experimental Psychology: Learning, memory and cognition*, 1982, *8*, 572-583.

Hatcher, C. & N. Robbins, *The development of reading skills in hearing-impaired children*. Ceder Falls: University of Northern Iowa, 1978.

Hirsch-Pasek, K., The metalinguistics of fingerspelling: An alternate way to increase reading vocabulary in congenitally deaf readers. *Reading Research Quarterly*, 1978, *22*, 455-475.

Hirsch-Pasek, K. & R. Treiman, Recoding in silent reading: can the deaf child translate print into a more manageable form? In: R.E. Kretchmer (Ed.), *Reading and the hearing impaired individual*, Washington, DC: The Alexander Graham Bell Association for the Deaf, 1982.

IJseldijk, F., De ontwikkeling van de woordenschat bij dove leerlingen. *Van horen zeggen*, 1989, *30*, 8-15.

Kleiman, G., Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior*, 1975, *4*, 323-339.

Kyle, J.G., Reading development of deaf children. *Journal of Research in Reading*, 1980, *3*, 86-97.

Leij, A. van der, Methodieken voor behandeling van ernstige leesproblemen. In: A. van der Leij & L.M. Stevens (Red.), *Dyslexie 1985*. Lisse: Swets & Zeitlinger, 1985.

Lichtenstein, E., *The relationships between reading processes and English skills of deaf college students*. Rochester, NY: National Technical Institute for the deaf, 1983.

Perfetti, C.A. & A.M. Lesgold, Discourse comprehension and sources of individual differences. In: P. Just & P.A. Carpenter, *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum, 1977.

Perfetti, C.A. & A.M. Lesgold, Coding and comprehension in skilled reading and implications for reading instruction. In: L.B. Resnick & P.A. Weaver, *Theory and practice of early reading* (vol. 1). Hillsdale, NJ: Erlbaum, 1979.

Reitsma, P., Word specific knowledge in beginning reading. *Journal of Research in Reading*, 1983, *6*, 41-56.

Reitsma, P., De verwerving van orthografische kennis. In: E. Assink & G. Verhoeven (red.), *Visies op spelling*. Groningen: Wolters Noordhoff, 1985.

Schaper, M.W. & P. Reitsma, Strategieën van dove kinderen in het onthouden van geschreven woorden. In: P. Reitsma, A.G. Bus & W.H.J. van Bon (red.), *Leren lezen en spellen*. Lisse: Swets & Zeitlinger, 1986.

Slowiczak, M.L. & C. Clifton, Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 573-582.

Stanovich, K.E., Individual differences in the cognitive processes of reading. *Journal of learning disabilities*, 1982, *15*, 485-497, 549-554.

Uden, A.M.J. van, Enkele psychologische aspecten bij het beginnend lezen van dove kinderen en hoe het verder gaat. *Speciaal Onderwijs*, 1982, *55*, 266-276.

Uden, A.M.J. van, Dove kinderen leren lezen, *Speciaal Onderwijs*, 1988, *61*, 82-90.

Walter, G.G., Lexical abilities of hearing and hearing-impaired children. *American Annals of the Deaf*, 1978, *123*, 976-982.

Wimister, S. & P. Arnold, The spelling strategies of deaf, partially hearing and hearing children. *Journal of the British Association of Teachers of the Deaf*, 1986, *10*, 146-152.

Curriculum vitae

M. W. Schaper studeerde psychologie (differentiatie Onderwijskunde) aan de Rijksuniversiteit Leiden. Zij is sedert 1971 als onderwijskundig medewerker verbonden aan Effatha, Christelijk Instituut voor doven.

Adres: Effatha, Zalkerbos 330, 2716 KS Zoetermeer

Manuscript aanvaard 5-6-'89

Summary

Schaper, M. 'The effect of articulation in memorising writting words by deaf children.' *Pedagogische Studiën*, 1990, *67*, 26-37

This paper centers on the question, how prelingually deaf children in the age of 6-8 years can best be instructed in the acquisition of orthographic vocabulary. In long standing experimental training 20 children weekly memorised a set of words in association with wordmeaning by procedures of practice-and-drill. Effects of two practice-conditions were evaluated. Half of the children memorised words without special instruction, the other part had to vocalise words during practice. These practice-conditions effectuated a dominant visual resp. articulatory strategy in memorising words, as was indicated by significant differences of relative difficulty in learning a set of visually versus articulatory low-discriminating words. Between both practice-conditions no significant differences were found in measures, directly involving the acquisition of orthographic vocabulary (amount of practice for reaching a fixed criterium for mastery, retention, orthographic knowledge). However, children who vocalised during training could memorise wordstrings significantly more effective, which indirectly may improve discourse-reading.

Boekbespreking

J. Zuidema, *Efficiënt spellingonderwijs. Een leer- en expertmodel voor het spellen*. Acco, Leuven/Amersfoort, 1988 X + 205 pag., *f*40.75, ISBN 90-334-18851.

In dit proefschrift wordt het probleem van de werkwoordspelling geanalyseerd en gezocht naar oplossingen. Omdat maatschappelijke weerstanden tegen spellingvereenvoudiging het niet waarschijnlijk maken dat deze op korte termijn doorgevoerd kan worden, wordt vooral aandacht besteed aan mogelijke verbeteringen van de didactiek. Daartoe worden eerst de spellingsregels behandeld, die aanleiding geven tot verwarring tussen verschillende mogelijke schrijfwijzen bij de werkwoorden. Vervolgens wordt nagegaan hoe in de loop der tijden geprobeerd is tot een optimale methode te komen voor het leren van de werkwoordspelling. Uit onderzoeksgegevens blijkt dat een algoritmische regelaanpak vermoedelijk een betere methode is dan de zogenaamde analogiemethode of de klassieke regelaanpak. Om een nieuwe stap te zetten op de weg naar verdere verbetering van de didactiek is een onderwijsexperiment uitgevoerd. Daarin werd geprobeerd om de effectiviteit van de instructie te verhogen. Twee factoren werden systematisch gevarieerd: de spelling*methode* en de *werkvorm*. Nagegaan werd het effect van de analogiemethode versus een algoritmische regelmethode en de klassikale versus de individuele werkvorm. De conclusie uit dit onderzoek is, dat de analogie-aanpak inderdaad te kort schiet en dat een regelgerichte algoritmische aanpak de voorkeur verdient. Opmerkelijk is dat deze regelgerichte methode vooral effectief is als er klassikaal les gegeven wordt in plaats van individueel.

In het derde hoofdstuk wordt nagegaan hoe fouten bij het spellen van werkwoorden beoordeeld behoren te worden. Er blijkt zich het paradoxale verschijnsel voor te doen dat het tijdelijk vóórkomen van bepaalde typen fouten een aanwijzing vormt voor een groeiende kennis van de werkwoordspelling. Het maken van fouten kan duiden op vorderingen van het leerverloop. De gebruikelijke werkwijze waarbij de voortgang wordt beoordeeld door het aantal gemaakte fouten op te tellen, schiet te kort. Het is mogelijk om positieve leerprocessen op het spoor te komen en te kwantificeren aan de hand van een gedetailleerde foutenanalyse. Daarom is een alternatieve beoordelingswijze ontworpen, die werd toegepast bij een heranalyse van de data van het bovengenoemd veldonderzoek.

Het eerder gevonden interactie-effect tussen de spellingmethode en de instructievorm blijkt dan nogrsterker te zijn dan eerder werd aangenomen. De conclusie dat een regelgeoriënteerde algoritmische methode de voorkeur verdient blijft gehandhaafd.

In theoretisch opzicht is vooral hoofdstuk vier interessant. Aan de hand van psychologische en taalkundige overwegingen wordt een zogenaamd expertmodel uitgewerkt. Een dergelijk model tracht de processen in kaart te brengen zoals die verlopen bij personen die de betreffende taak zodanig beheersen dat die processen feilloos en zoveel mogelijk geautomatiseerd verlopen.

De auteur is zich heel goed bewust hoe moeilijk het is een dergelijk model te ontwikkelen. Hij stelt zich dan ook voorzichtig op. De pretenties van het gepresenteerde model zijn bescheiden: 'het moet gezien worden als een eerste aanzet voor het in kaart brengen van EXPERTdenken. Er is niet gestreefd naar volledigheid.' Wel pretendeert het model datgene te kunnen verklaren wat bekend is met betrekking tot de problematiek van het 'woordbeeld'. Bij de bespreking van het model vindt een zekere controle plaats aan de hand van voorbeelden. Uitdrukkelijk wordt opgemerkt dat het model daardoor aannemelijker wordt gemaakt maar niet bewezen. Het is niet mogelijk het model in het kort weer te geven. Daarvoor is het te complex. Een belangrijk aspect verdient echter aandacht. Bij het ontwerpen van het model is als uitgangspunt gekozen een 'psychologische economie'. Dit houdt in dat bepaalde processen zoveel mogelijk geautomatiseerd verlopen. Het model maakt gebruik van zogenaamde 'filters'. Het idee daarbij is, dat gemakkelijke spellingsgevallen automatisch uitgefilterd worden op basis van patroonherkenning. Door sommige patronen ligt direct een schrijfwijze vast. Maar bij andere patronen is dat niet het geval. Alleen als patronen geen zekerheid geven over de correcte schrijfwijze wordt een 'alarm' ingeschakeld. Als een 'alarm' ingeschakeld is, kan de correcte schrijfwijze gevonden worden na een nadere analyse van het probleem. De metafoor 'filter' moet in dit model niet verstaan worden als iets passiefs, maar veronderstelt een constante activiteit van de speller op allerlei niveaus van taalverwerking.

Belangrijk is verder dat onderscheid gemaakt wordt tussen een algoritme voor de *introductie* van de werkwoordspelling en een algoritme van de *gevorderde* spellers. Beide algoritmen worden gepresenteerd als discussiestof en nodigen uit tot kritiek.

Pedagogische Studiën 1990 (67) 38-41

Een punt van kritiek zou kunnen zijn, dat in de modellen alleen cognitieve variabelen zijn opgenomen, terwijl bekend is dat bij leerprocessen ook niet-cognitieve variabelen een rol spelen. Als men het op prijs stelt feilloos te spellen, dan zou dit wel eens een positieve invloed kunnen hebben op het activeren van filterprocessen. Het gegeven dat studenten meestal niet uitgelegd behoeft te worden waarom een werkwoordsvorm in hun werkstuk of scriptie foutief geschreven is, zou een aanwijzing kunnen zijn dat motivationele aspecten verantwoordelijk kunnen zijn voor spelfouten.

Als toegift wordt een onorthodox voorstel gedaan voorrspellingvereenvoudiging, waarbij rekening wordt gehouden met het feit dat spelling niet alleen van belang is voor schrijven, maar ook voor lezen.

Onderwijskundigen, moedertaaldocenten en onderzoekers vinden in dit proefschrift een goed overzicht van de stand van zaken op dit terrein en meer dan voldoende stof voor nadere reflectie en onderzoek.

M. J. C. Mommers

J. Vos, *Meer dan feiten ... Werken met het informatieve kinderboek op de basisschool.* Uitgeverij De Ruiter, Gorinchem, 1989, 136 pag., *f*37,50, ISBN 90-05-161043.

Kinder- en jeugdboeken zijn niet meer weg te denken bij het moedertaal-onderwijs. Het aantal boeken dat richtlijnen geeft voor het gebruik van kinder- en jeugdboeken breidt zich steeds verder uit. Jacques Vos neemt een groot deel van deze boeken voor zijn rekening (zie Pedagogische Studiën, 1988, p. 422-423).

Langzamerhand verovert het kinderboek een plaats in het gehele onderwijs. In "Meer dan feiten ..." schenkt Vos aandacht aan een boeksoort die tot nu toe in het verdomhoekje heeft gezeten, namelijk de informatieve kinder- en jeugdliteratuur. Het boek is in eerste instantie bestemd voor leerkrachten in het basisonderwijs en de onderbouw van het voortgezet onderwijs. Ook voor de opleidingen zou het te gebruiken zijn.

Het boek bestaat uit twee gedeeltes. In deel 1 wordt achtergrondinformatie over het informatieve kinderboek gegeven; in deel 2 wordt aangegeven hoe een en ander in de praktijk uitgewerkt kan worden.

In het gedeelte over de *achtergrond*informatie (hoofdstuk 1 t/m 4) geeft de auteur na een lange inleiding een definitie van het informatieve kinderboek. "Een informatief boek verschaft de lezer informatie over zaken, gebeurtenissen en verbanden op zo'n manier, dat door een specifiek taalgebruik en structurele middelen de lezer tegelijkertijd iets leert en zich vermaakt". Deze definitie ontleent hij aan 'een boek van de Duitse jeugdliteratuurdeskundige Klaus Doderer'. Jammer genoeg ontbreekt hier, en het begint langzamerhand traditie te worden bij Vos, een literatuurverwijzing. Vervolgens gaat hij in op de indeling in verschillende soorten informatieve boeken, allereerst een indeling naar 'de themavelden voor wereldoriëntatie' en vervolgens een indeling die verder in het boek gebruikt zal worden, naar functie. Informatieve boeken kunnen als functie hebben de kinderen bepaalde begrippen bij te brengen, kinderen aan te zetten tot het zelf uitvoeren van proeven e.d., kinderen alle mogelijke aspecten van bepaalde onderwerpen laten zien of kinderen op een snelle manier informatie verschaffen.

In hoofdstuk 2 bespreekt Vos aan de hand van een aantal voorbeelden de relatie tussen de doelstelling van de auteur van een informatief boek en de uiteindelijke vorm waarin het aangeboden wordt aan kinderen. Hoofdstuk 3 heeft betrekking op het beoordelen van boeken. De volgende aandachtsvelden worden besproken: 'begrijpelijk schrijven' (taalgebruik), betrouwbaarheid van de informatie, de vormgeving en de bruikbaarheid voor hetronderwijs. Een en ander wordt schematisch weergegeven. In het laatste hoofdstuk van deel 1 worden enkele boeken over hetzelfde onderwerp met elkaar vergeleken. De criteria uit hoofdstuk 3 worden hiervoor gebruikt.

Deel 2 heeft als titel: *didactiek en methodiek*. In dit deel wordt uitgebreid ingegaan op het gebruik van informatieve kinderboeken in het onderwijs. Alvorens echter met deze boeken aan de slag te kunnen, zullen de leerlingen een aantal algemene vaardigheden moeten bezitten (begrijpend lezen, het kunnen alfabetiseren, interpreteren van schema's etc.). In hoofdstuk 6 (vier stappen) baseert Vos zich op het werk van Theodor Karst. Deze onderscheidt bij het werken met boeken vier fasen: voorinformatie, analyse van de inhoud, didactische analyse en aanpak in de klas. In praktische voorbeelden kleedt Vos dit schema aan. In hoofdstuk 7 worden allerlei deelvaardigheden, zoals het gericht zoeken van informatie, de index etc. besproken. Opvallend is dat Vos, terecht, specifiek aandacht besteedt aan het 'illustraties lezen'. In hoofdstuk 8 t/m 11 komen achtereenvolgens praktische lessuggesties met betrekking tot 'toepassen', 'kritisch lezen', 'boekpromotie' en 'boeken bij een project' aan de orde.

"Meer dan feiten ..." is allereerst een zeer praktisch boek. Veel lessuggesties kunnen zo onder de copieerder worden gelegd en worden uitgeprobeerd. Gezien de doelgroep zeker een verdienste.

Het eerste deel is goed uitgewerkt, al is er met betrekking tot de achtergrondinformatie sprake van veel overlap. Op blz. 34 stelt Vos dat een informatief boek toegankelijker wordt als er "een samenvatting aan het einde van het hoofdstuk en/of eventueel een soort leeswijzer aan het begin" zou zijn opgenomen. Vos heeft dit zelf ook gedaan en in een hoofdstuk van nog geen tien pagina's lees je driemaal hetzelfde. Dit is mij teveel. De schematische weergave daarentegen aan het eind van het derde hoofdstuk (blz. 39-40) over het beoordelen van boeken is zeer handig en bruikbaar. Vaker is er overlap. Op blz. 54 lees ik voor de tiende maal dat we "informatieve boeken zeker niet mogen (ver)didactiseren". Mensen die het werk van Jacques Vos kennen, zullen zijn publikaties uit "Leestekens" en "Project jeugdliteratuur" herkennen.

Het boek levert, zoals gezegd, een praktische handleiding om informatieve boeken te gebruiken in het onderwijs. Voor leerkrachten basisonderwijs een welkome aanvulling op de reeds bestaande didactiekboeken. Voor de opleidingen is het boek ook geschikt. Hebben de mensen in de praktijk in het algemeen geen tijd om zich door achtergronden heen naar de praktische suggesties te worstelen, voor studenten PABO is dit uitermate geschikt. Het eerste deel levert ook vaak discussiestof. Zo signaleert Vos 'het probleem van het antropomorfiseren' van kinderboeken. In hoeverre dit is toegestaan, wordt niet uitgewerkt. Dit geldt mijns inziens ook voor de stelling dat er een zeer nauwe relatie dient te bestaan tussen 'doelstellingen, visiervan de auteur(s) en vorm' van een informatief boek.

Het is belangrijk dat er eindelijk een boek gewijd is aan het informatieve kinderboek. "Meer dan feiten ..." is voor mensen in de praktijk en voor de toekomstige mensen in de praktijk geschikt.

Bij het ontbreken van bronvermeldingen zullen we ons maar neerleggen. In het voorwoord schrijft Vos "Af en toe heb ik mij ook voorzichtig opgesteld; ik hoop dat de lezer voor lief neemt dat mijn tekst daardoor soms geen uitgesproken karakter vertoont". Een dergelijke opmerking maakt een bespreking van het boek niet eenvoudiger, want een recensent moet een uitgesproken mening over het boek geven.

A. W. M. Duijx

F. Goffree & K. Buys, *Tegengesteld, wiskundedidactiek op de grens van basis- en voortgezet onderwijs, toegespitst op negatieve getallen*, Bekadidact, Baarn, 1989, 247 pp. *f* 34,50, ISBN 90 321 0634 1.

Deze bundel kwam tot stand dankzij de samenwerking tussen een achttal auteurs, vrijwel allen werkzaam aan een van de Nederlandse lerarenopleidingen. Wanneer auteurs besluiten samen een boek te schrijven en de lezers vervolgens tracteren op een bundel waarin ze elkaar bij herhaling tegenspreken, is er normaal gesproken iets misgegaan. De redactie miste op cruciale momenten het overzicht of de tegenspraken werden op onbewaakte momenten over het hoofd gezien. Er kan echter ook iets heel anders achter zitten, namelijk dat de redactie juist op die tegenspraken uit was.

En inderdaad is dit het geval. De auteurs van dit boek prikkelen elkaar telkens tot kritische reacties. Deze werkwijze leverde geen bundel op, waarin zekerheden worden geëtaleerd maar – en dat is vaak veel interessanter – een bundel waarin de lezer wordt geconfronteerd met vragen, vermoedens en suggesties die hem of haar soms in onzekerheid laten maar daardoor wellicht stimuleren tot eigen oordeelsvorming.

Het boek gaat over negatieve getallen (n.g.) in de brugperiode, dus bij de overgang tussen basisonderwijs en voortgezet onderwijs. Negatieve getallen kunnen beschouwd worden als een uitbreiding van de verzameling van natuurlijke getallen en wel zodanig dat er sprake is van permanentie van rekenregels. Dat wil zeggen dat de 'oude' regels (bijvoorbeeld distributiviteit) ook gelden voor het 'nieuwe' gebied van de n.g.. Dit is in essentie, en uiterst beknopt, de wiskundige achtergrond van de nu besproken publikatie. In deze recensie beperk ik me tot de, mijns inziens, meest relevante onderwijskundige vraagstukken die in "Tegengesteld" worden behandeld.

Het boek begint met een exploratieve analyse van het begrip dat kinderen uit de achtste groep van n.g. hebben. De kinderen blijken allerlei ideeën over n.g. te hebben, soms zijn die bruikbaar, maar vaak ook misplaatst ('Ze bestaan niet, want –7 knikkers dat bestaat niet.'). In het tweede hoofdstuk worden de kinderen vanuit de realiteit geconfronteerd met n.g. en dat experiment verloopt allerminst van een leien dakje. In het derde hoofdstuk wordt verslag gedaan van confronterende gesprekken over n.g. maar nu met brugklassers. Na een jaar onderwijs werden dezelfde kinderen nogmaals aan de tand gevoeld omtrent hun kennis van n.g.. Voor sommige kinderen leiden die getallen nog steeds een duister bestaan en buiten de wiskunde komen ze, zo denken sommigen, niet voor. Wat zeggen de school-

boeken over n.g.? In het vierde hoofdstuk wordt op die vraag antwoord gegeven. De ideeën van schoolboekenauteurs zijn zeer uiteenlopend, van: n.g. zijn er gewoon tót: je kunt ze zelf maken. De auteurs pleiten, uitgaande van een zogenoemde realistische kijk op wiskundeonderwijs zoals de lezers al wel begrepen zullen hebben, voor de laatste aanpak. N.g. moeten, zo stellen ze, niet louter formeel – in formules dus – behandeld worden. Hoofdstuk vijf moet iedereen lezen die na het voorafgaande zou denken dat n.g. met enkele voorbeelden makkelijk uit te leggen zijn. De auteur van dit hoofdstuk maakt glashelder hoe moeilijk n.g. voor kinderen zijn maar ook nog voor volwassenen die gewend zijn vooral vanuit formele kaders (zeg maar: blinde regels) te werken. Vervolgens worden n.g. beschouwd vanuit het oogpunt van een longitudinale planning. Opmerkelijk is het dat, blijkens interviewgegevens, volwassenen vaak bitter weinig van het onderwijs in n.g. hebben opgestoken.

En nu in didactisch opzicht een kernprobleem, hoe leg je uit: $- \times - = +$? In de schoolboeken werden acht aanpakken aangetroffen, waaronder het verstrekken van de definitie, geconstrueerde fysische modellen en deductieve redeneringen. Al deze benaderingen zijn voor de kinderen erg lastig en staan ver van hun realiteit. Veel lezers vragen zich – wellicht ietwat ongeduldig – af hoe het nu eigenlijk moet met die n.g.. Pas in het negende hoofdstuk worden de 'didactische dilemma's' besproken. Een eerste dilemma luidt of we n.g. direct in de wiskunde zelf moeten zoeken of eerst in contexten. Een eenduidig antwoord lijkt onmogelijk want als we n.g. moeten optellen zijn contexten makkelijk te vinden, maar bij vermenigvuldigen en delen neemt de zingeving door middel van contexten snel af. Moeten in die gevallen dan maar, en zonder omhaal, wiskundige regels gepresenteerd worden? Dat is weer strijdig met het uitgangspunt dat in de wiskunde regels altijd plausibel gemaakt moeten worden. Maar hoe bereiken de kinderen dan dat formele niveau? In het dertiende hoofdstuk worden de contouren geschetst van een 'negatieve getallen-didactiek'. Waarom uitgaan van contexten, zo luidt een belangrijke vraag. Doordat het kind zich 'beelden' vormt (denk bijvoorbeeld aan inkomsten en uitgaven) ontstaan in het denken referentiepunten waarop een kind zich kan oriënteren bij het oplossen van problemen binnen dat onderwerp. Maar tegelijkertijd moeten we uitkijken voor het gevaar van afdwalen in voor de wiskunde onproduktieve richtingen. Maar niet alleen de context is van belang, de leraar moet ook idee hebben van de intuïtieve noties van de kinderen. Wat denkt een kind bijvoorbeeld bij aftrekken 'wat er toch al niet is'?

Waar het in essentie om gaat, is dat de kinderen zinvolle voorstellingen kunnen ontwikkelen vanuit contexten. En dit laatste betekent dat contexten aanleiding moeten geven tot heruitvinden van wiskundige modellen, structuren en wetmatigheden. Vaak echter zal de wiskunde *zelf* context moeten zijn. Bij de overgang van het basis- naar het voortgezet onderwijs gaat de wiskunde als context een steeds grotere rol spelen. In plaats van werken in het systeem ontstaat nu een denken over het systeem.

In "Tegengesteld" wordt een scherp beeld gegeven van de problemen waar een leraar, die in zijn/haar klas een onderwerp als n.g. behandelt, voor komt te staan. De leerlingen begrijpen er niets van en de leraar begrijpt niet hoe het komt dat de leerlingen het niet begrijpen. En recepten zijn voor dit probleem niet voorhanden, zelfs in de wiskundige vakdidactiek weet men er niet goed raad mee. Het gebied van de n.g. lijkt zich te onttrekken aan de didactische theorie en regels die op vele andere gebieden wél bruikbaar bleken. Maar hoe vaak wordt niet gedacht dat het in de wiskunde allemaal zo klaar is als een klontje? Wiskunde is exact en dus overzichtelijk en ook didactisch kent dit vak geen geheimen meer. Wie dit boek leest, is van deze illusie spoedig verlost en alleen daarom al kan ik het van harte aanbevelen.

J. M. C. Nelissen

Mededelingen

Geldpremies voor overzichtsartikelen alfa- en gammawetenschappen

Auteurs van overzichtsartikelen op het terrein van de geesteswetenschappen en van de gedrags- en maatschappijwetenschappen kunnen ook in 1990 in aanmerking komen voor een geldpremie.

Voor de bekroningen is er dit jaar voor elk der beide vakgebieden een bedrag van *f* 1.000 à *f* 5.000. De toekenning gebeurt door de Commissie Geesteswetenschappen en de SociaalWetenschappelijke Raad van de Koninklijke Nederlandse Akademie van Wetenschappen.

Onder geesteswetenschappen worden verstaan: taal- en letterkunde, historische wetenschappen, wijsbegeerte en godgeleerdheid.

Het begrip gedrags- en maatschappijwetenschappen wordt ruim opgevat: ook rechtswetenschap en economie vallen eronder.

Om voor bekroning in aanmerking te komen moet het artikel in 1989 of 1990 zijn verschenen in een Nederlands wetenschappelijk tijdschrift.

De volgende maatstaven worden aangelegd:

1 Het artikel moet inhoudelijk een aanmerkelijk segment van een erkende discipline bestrijken.
2 De hoofdlijnen van de ontwikkelingen in een bepaald tijdvak moeten erin worden beschreven, zonder dat er relevante literatuur is overgeslagen.
3 De auteur dient eigen uitgangspunten op methodologisch gebied en andere theoretische veronderstellingen van principiële aard ten minste impliciet duidelijk te maken.
4 Er dient te worden ingegaan op het theoretisch gehalte van de ontwikkelingen in (het segment van) de discipline.

Auteurs die voor bekroning in aanmerking wensen te komen, moeten hun artikel vóór 1 juli 1990 inzenden aan: Commissie Geesteswetenschappen, resp. Sociaal-Wetenschappelijke Raad, Herengracht 410-412, 1017 BX Amsterdam.

PAOS Cursussen 1990

Onlangs is verschenen 'Cursussen 1990', de brochure met het programma voor 1990 van het Postacademisch Onderwijs Sociale Wetenschappen (PAOS). De brochure bevat aankondigingen van ruim veertig cursussen, congressen en symposia op het gebied van psychologie, pedagogiek, psychoterapie, diagnostiek, ontwikkeling en onderwijs. De brochure is kosteloos aan te vragen bij: PAOS, Postbus 325, 2300 AH Leiden, tel.: 071 – 143141.

Confrontatie van culturen: Onderwijsvernieuwing in de hedendaagse praktijk.

De Mennicke Stichting voor Opvoeding en Onderwijs, het Samenwerkingsverband van Organisaties voor Onderwijsvernieuwing en het Post Academisch Onderwijs Sociale Wetenschappen organiseren op 14 februari 1990 te Utrecht een studiedag met als thema: Confrontatie van culturen: Onderwijsvernieuwing in de hedendaagse praktijk.

Voor inlichtingen en inschrijving: PAOS, Postbus 325, 2300 AH Leiden, tel.: 071 - 143141.

Congres 'Pedagogiek in beweging'

De Vereniging tot Bevordering van de Studie der Pedagogiek (VBSP) zal ter gelegenheid van haar 65-jarig bestaan op zaterdag 10 februari 1990 een congres houden met als thema 'Pedagogiek in beweging'. Dit congres wordt gehouden in de gebouwen van de Rijksuniversiteit te Utrecht.

Belangstellenden, ook (oud)studenten, kunnen zich voor nadere inlichtingen wenden tot het secretariaat van de VBSP, Herengracht 266, 1016 BV Amsterdam, tel. 020 - 225496.

Symposium Vereniging Pedagogen voor de Vrede

De Vereniging Pedagogen voor de Vrede houdt op 17 februari 1990 te Utrecht een symposium met als titel 'Door de buis genomen? De invloed van televisiegeweld op kinderen'.

Nadere inlichtingen zijn verkrijgbaar bij: S. Miedema, tel.: 071 - 273447, of G. Jacobs, tel.: 071 - 273455; Rijksuniversiteit Leiden.

WACRA '90 Conferentie

In samenwerking met de World Association for Case Method Research and Application (WACRA) en The European Network for System Simulation and Management Gaming (Deutsche Planspiel Zentrale) organiseert de Universiteit van Twente van 9 tot 11 april 1990 een conferentie over 'Problem Solving with Cases'.

Voor nadere informatie kunt u contact opnemen met mevrouw J. Oosterhuis, Universiteit Twente, Postbus 217, 7500 AE Enschede, tel.: 053 - 356695.

Oprichting Vereniging RAIN

Recentelijk is de 'Reading Association In the Netherlands' (RAIN) opgericht, een Nederlandse afdeling van de 'International Reading Association' (IRA). De vereniging stelt zich o.a. als doel: Het stimuleren van onderzoek naar leesproblemen en leeronderwijs, de verspreiding van informatie op het terrein van leesonderzoek, instrument- en leerplanontwikkeling en beleid op het gebied van lezen en leesonderwijs.

Nadere inlichtingen zijn verkrijgbaar bij: Drs. J.H.M.W. Clemens, Hogeschool Midden Nederland, Postbus 14007, 3508 SB Utrecht.

Symposium Begrijpend Lezen

RAIN organiseert in juni 1990 een symposium over 'Instructie in begrijpend lezen'. Een ieder die geïnteresseerd is op dit symposium een bijdrage te leveren wordt verzocht een voorstel in te dienen vóór 12 februari 1990.

Nadere informatie over de inhoud van het symposium en richtlijnen voor papers zijn verkrijgbaar bij: Drs. A.M.A. Walraven, Paedologisch Instituut, Postbus 303, 1115 ZG Duivendrecht.

Conference High Ability over the Life Span, Utrecht, June 21-23, 1990

Recently there has been a growing concern with the study and development of high ability and high level performance especially in childhood and adolescence. Studies about maintaining high level performance over the life span, however, are still scarce.
The main topic of this conference is the way in which high ability individuals succeed in performing at high levels throughout the whole life span, in view of the continuous changes that take place in cognition and personality and in view of the increased speed of social and technological change.

Keynote speakers will be: Dr. Dean Simonton (University of California), Dr. Howard Gardner (Harvard University), Dr. Patric Rabbitt (University of Manchester, U.K.), Dr. Georg Rudinger (University of Bonn) and Dr. Ger Boerlijst (University of Twente).
The conference will start with a forumdiscussion. In this forum will participate well known people who are succesful in maintaining a high level of performance in their respective fields of interest.

For information: Dr. W.A.M. Kok (secr.), ISOR, Rijksuniversiteit Utrecht, Heidelberglaan 2, 3584 CS Utrecht, tel.: 030 - 534867/534555.

Inhoud andere tijdschriften

Comenius
9e jaargang, nr. 3, 1989

De geest uit de machine? 'Nieuwe media' in het onderwijs, door W. Wardekker
Pleidooi voor een 'kies'keuriger onderwijsemancipatiebeleid, door M. Urlings en M. Volman
'Individualpsychology' en onderwijs. Opvattingen en initiatieven van drie generaties Adlerianen, door E. Vanpoucke
Jongeren en hun kijk op ouderschap, door F. van Wel
Ze komen nooit op tijd! Het cultuurverschil als verklaring voor ongeregeldheden, door J. van der Zwaard
Geen scholing maar ontwikkeling. Pleidooi voor een pedagogisering van het onderwijs, door W. Pols
De naïeve bril van Dronkers, door S. Soutendijk
Weerwoord op Soutendijks 'De naïeve bril van Dronkers', door J. Dronkers.

Comenius
9e jaargang, nr. 4, 1989

Pedagogiek controversieel. Inleiding op het thema, door J.D. Imelman en A. Rang
Eugenetica en sociale arbeid. Een Nederlandse discussie over bevolkingskwaliteit tijdens het Interbellum, door J. Noordman
Volksopvoeding, nationalisme en cultuur. Nederlandse denkbeelden over massa-educatie in het Interbellum, door W. Krul
Kunst en pedagogiek: een controversiële relatie, door M. van der Kamp
Überlegungen zum Verhältnis von Pädagogik und Avantgarde in der Weimar Republik, door A. Rang
Rassentheoretische Vorstellungen und ihre Bedeutung für den humanistischen Unterricht (1925-1936), door S. Bittner
Pädagogisches Handeln im Spannungsfeld zwischen Kaiserreich und Drittem Reich. Das Gymnasium und seine Lehrer in den Rheinlanden zwichten 1919 und 1936, door H-J. Apel

Pedagogisch Tijdschrift
14e jaargang, nr. 5, 1989

Onderwijsbeleid: ontwikkelingslijnen en vooronderstellingen van een studiedomein, door J.L. Vanderhoeven
Herzieningen in het onderwijsbestuur, door A.M.L. van Wieringen
Autonomievergroting en het benutten van de beleidsruimte door scholen, door P. Sleegers en A. van Rooyen
De plaats van de provincie in het Nederlandse onderwijsbestel en beleid, door S. Karsten en C.C. Verhoeff

Invloed van decentralisatie en deregulering op het schoolbeleid, door J.C. Verhoeven en R. Vandenberghe
Mogelijkheden van beleidsevaluatie bij grootschalige onderwijsvernieuwing, door P.N. Karstanje
Ouders, school en participatie, door T. Pittoors en J.L. Vanderhoeven

Tijdschrift voor Onderwijsresearch
14e jaargang, nr. 4, 1989

Computergeleid enquêteren: een overzicht van technologische vernieuwingen, door E.D. de Leeuw
De effecten van een opleiding die tot doel heeft a.s. leraren te leren reflecteren, door F.J.A. Korthagen en Th. Wubbels
Effecten van tweetalig onderwijs aan Marokkaanse kleuters, door E. Wagenaar en E.M. Scholte
Onderwijs en maatschappelijke ongelijkheid. Een overzicht van 25 jaar theorievorming en onderzoek in Nederland, door J.C.C. Rupp en A.A. Wesselingh
Notities en Commentaren
A nonlinear correction on linear equating, door D.N.M. de Gruijter

Tijdschrift voor Orthopedagogiek
28e jaargang, nr. 10, 1989

Ik kom er wel, door H. van Tilburg en H. Bleeker
Vroegtijdige onderkenning van taalontwikkelingsstoornissen, door S.M. Goorhuis-Brouwer
Het interview met behulp van anatomisch correcte poppen, door H. Baartman, S.M.J. van Hekken en S. Bierens

Tijdschrift voor Orthopedagogiek
28e jaargang, nr. 11, 1989

Rispens' problematische opvoedingstheorie, door J.J.C. van Leeuwen
Is de terugkeer van Rispens een vooruitgang? door P.A. de Ruyter
Apartheidspolitiek in orthopedagogenland, door J. Van Acker
Over het object van de orthopedagogiek, door J.W. Steutel
Het bevorderen van het gebruik van leesstrategieën via een dialogische interventieprocedure bij kinderen met leerproblemen, door A. Pennings, T. de Jong en M. Laudy

Ontvangen boeken

Asselbergs-Neessen, V., *Kind, kunst en opvoeding*. Acco, Amersfoort, 1989, *f* 49,-.

Creemers, B., T. Peters, D. Reynolds (Eds.), *School effectiveness and school improvement*. Swets & Zeitlinger, Lisse, 1989, *f* 72,50.

Gemert, G.A. van, Th.A.H. Deenen, J.T. Timmer (red.), *Psychologen en pedagogen in de zwakzinnigenzorg*. Stichting Kinderstudies, Groningen, 1989.

Gilst, J.C. van, J. Kugel, K.G. van der Straten, *Speciale bewegingsopvoeding*. Intro, Nijkerk, 1989, *f* 39,50.

Hagen, A.M.P. & H. Knoors (red.), *Onderwijs aan doven*. Swets & Zeitlinger, Lisse, 1989, *f* 39,50.

Hoksbergen, R & L. Brunjes, *Een buitenlands kind adopteren*. Gids voor adoptie-ouders en hun raadgevers. Acco, Amersfoort, 1989, *f* 19,90.

Hubbard, F., *Ingaan op huilen bij baby's*. Een kritische replicatie van de Baltimore huilstudie. Rijksuniversiteit Leiden, 1989.

Lentz, L.R. & H. van Tuyl, *Het leerplan in Nederland*. Een studie naar vorm, inhoud en functie van het leerplan voor het primair onderwijs (dissertatie). Instituut voor Leerplanontwikkeling (SLO), Enschede, 1989.

Linden, W.J. van der, L.J.Th. van der Kamp (red.), *Meetmethoden en data-analyse*. Swets & Zeitlinger, Lisse, 1989, *f* 39,90.

Mahieu, P., *Mensen, macht en management*. Een onderzoek naar schoolorganisaties. Acco, Amersfoort, 1989, *f* 36,50.

Michielse, H.C.M. & F.J. Berkers, *Welzijn en discipline*. Van tuchthuis tot psychotherapie. Boom, Meppel, 1989, *f* 38,50.

Mooij, A.M., F. Koenraadt, J.M.J. Lommen-van Alphen (red.), *De persoon van de verdachte*. De rapportage pro justitia vanuit het Pieter Baan Centrum. Gouda Quint, Arnhem, 1989, *f* 19,-.

Mooren, P. & H. Verdaasdonk (red.), *Kinderen en non-fictie*. Zwijsen, Tilburg, 1989, *f* 39,-.

Rink, J.E., R.C. Vos, H.M. van Lokven, C.M. Slagveer, *Grensgevallen*. De houding van jongeren t.o.v. sociale grenzen. Acco, Amersfoort, 1989, *f* 32,50.

Scheerens, J. (red.), *Evaluatie: om de kwaliteit van het onderwijs*. Swets & Zeitlinger, Lisse, 1989, *f* 19,90.

Slavenburg, J.H. & T.A. Peters (red.), *Het project Onderwijs en Sociaal Milieu: een eindbalans*. Rotterdamse School Advies Dienst, Rotterdam, 1989, *f* 35,-.

Swaans-Joha, B.C. & J.J. Hox (red.), *Praktijkgericht onderzoek*. Debat, design, data. Acco, Amersfoort, 1989, *f* 29,-.

Tomic, W. & P.C. van der Sijde, *Changing teaching for beter learning*. Swets & Zeitlinger, Lisse, 1989, *f* 29,95.

Treffers, A., E. de Moor, E. Feijs, *Proeve van een nationaal programma voor het rekenwiskundeonderwijs op de basisschool*. Zwijsen, Tilburg, 1989, *f* 27,50.

Tromp, Th.J.M., *The acquisition of expertise in computer programming*. Uitgeverij Thesis, Amsterdam, 1989, *f* 49,50.

Omgaan met combinatieklassen: een onderzoek naar de effecten van een teamgericht nascholingsprogramma*

S. VEENMAN, P. LEM en E. ROELOFS**
Facultaire Vakgroep Interdisciplinaire Onderwijskunde
Katholieke Universiteit Nijmegen

Samenvatting

Steeds meer basisscholen in Nederland worden gedwongen om combinatieklassen samen te stellen. De meeste leerkrachten zijn nauwelijks voorbereid op deze in veel gevallen ingewikkelde klasse-organisatie. Een quasi-experimenteel onderzoek (experimentele groep – controlegroep) is uitgevoerd om de effecten te meten van het teamgerichte nascholingsprogramma 'Omgaan met Combinatieklassen'. Gegevens uit onderzoek naar combinatieklassen, effectieve instructie en klassemanagement zijn vertaald in leerkrachtgedragingen. In totaal hebben 41 leerkrachten afkomstig uit 8 schoolteams 7 cursusbijeenkomsten van 3 uur bijgewoond. Voorafgaande en na afloop van de cursus zijn bij 26 leerkrachten (experimentele groep N= 17; controlegroep N= 9) observaties uitgevoerd. Uit de gegevens blijkt dat een significant cursuseffect optreedt ten aanzien van het taakgerichte gedrag van de leerlingen en ten aanzien van een aantal leerkrachtgedragingen: effectieve instructie, klassemanagement en -organisatie.

* Het onderzoek 'Omgaan met Combinatieklassen' is uitgevoerd met financiële steun van het Instituut voor Onderzoek van het Onderwijs (SVO-project 6603).

** De auteurs danken voor hun bijdrage aan het onderzoek: F. Nijssen, H. Lassche, M. Stadelman en C. Veraart en alle schoolteams, Pabodocenten en schoolbegeleiders die aan het onderzoek hebben deelgenomen.

1 *Achtergrond*

In toenemende mate worden basisscholen in Nederland gedwongen om groepen met leerlingen van verschillende leeftijd, ontwikkelings- en kennisniveau samen te voegen in één klas of groep. Een steeds groter aantal leerkrachten is, voornamelijk als gevolg van het sterk teruglopend leerlingenaantal in de afgelopen vijftien jaar, voor deze zogenaamde combinatieklassen geplaatst. In 1989 bestaat naar schatting 30 tot 40 procent van alle klassen uit gecombineerde klassen. In tegenstelling tot vijftien jaar geleden – combinatieklassen blijken dan voornamelijk voor te komen in kleine scholen op het platteland – komen scholen met combinatieklassen nu ook in groten getale voor in de steden. Scholen die noodgedwongen moeten overgaan tot combinatieklassen en die geen ervaring hebben met deze organisatievorm, stellen aan hun leerkrachten hoge eisen voor wat betreft een doelmatige inzet van groeperingsvormen, leermiddelen en instructietijd. Mede door de bijkomende taakbelastingen ervaren veel leerkrachten hieromtrent problemen (Lee, 1984).

In Nederland is het onderzoek naar de combinatieklas betrekkelijk laat op gang gekomen. De aandacht is lange tijd opgeëist door onderzoek naar het bestaansrecht en de gevolgen van dreigende sluiting van kleine scholen (Santema, 1974). Ook in de landen om ons heen is in de jaren zestig discussie gevoerd over de kleine plattelandsschool, zoals in de Bondsrepubliek, het Verenigd Koninkrijk en Scandinavië (Veenman, Lem, Voeten, Winkelmolen & Lassche, 1986). In deze context staat het onderwijs in combinatieklassen nog nauwelijks ter discussie. Einde jaren zeventig publiceert de Inspectie Basisonderwijs (1978) in het 'Onderwijsverslag' de resultaten van een enquête onder leerkrachten over de problematiek van de combinatieklas. Uit deze enquête blijkt dat veel leerkrachten de combinatieklas als nadelig voor het onderwijs beschouwen. De bezwaren die worden genoemd zijn velerlei: de leerlingen in combinatieklassen krijgen minder individuele aandacht;

de zwakke en betere leerlingen komen te kort; voor de zwakste leerlingen is de combinatieklas het meest nadelig; de leerlingen worden over en weer afgeleid; het werken in combinatieklassen is veel zwaarder en vraagt meer tijd aan voorbereiding en correctie; degene die een combinatieklas heeft, moet vaak als eenling 'anders' werken en vindt weinig steun bij meer bevoorrechte collega's met gewone klassen; de methoden die op school worden gebruikt, zijn ingesteld op de jaarklas; er is minder gelegenheid tot mondeling actief taalgebruik, onder andere in verband met de instructie aan meer dan één klas. Ook veel ouders blijken de combinatieklas doorgaans te zien als nadelig voor hun kinderen. Ze zijn van mening dat de leerlingen in de combinatieklas minder individuele aandacht krijgen, te gauw zijn afgeleid en zich minder goed kunnen concentreren. Men vreest achterraken op leerlingen in 'gewone' klassen en uiteindelijk moeilijkheden bij de overgang naar het voortgezet onderwijs.

Dit onderwijsverslag rapporteert voor het eerst opvattingen over vermeende voor- en nadelen. Gegevens uit empirische studies naar het functioneren van combinatieklassen zijn op dat moment nog niet voorhanden.

Tussen 1981 en 1985 zijn in combinatieklassen in Nederlandse basisscholen drie observatiestudies uitgevoerd naar het leren van de leerlingen en de instructie-activiteiten van de leerkrachten: twee vooronderzoeken (Veenman, Lem & Winkelmolen, 1985) en een hoofdonderzoek (SVO-project 1059) (Veenman e.a., 1986). In de voornoemde nadelen van de combinatieklas speelt de factor tijd een belangrijke rol. De resultaten uit onze studies laten zien dat tussen combinatieklassen onderling grote verschillen bestaan in het gebruik van de leertijd door de leerlingen en in de toegepaste managementvaardigheden van de leerkrachten. In de onderzochte klassen ligt het taakgerichte gedrag van de leerlingen in combinatieklassen ongeveer 6% lager dan in enkelvoudige klassen. Deze taakgerichte leertijd wordt beïnvloed door het gekozen groepsarrangement waarin de leerlingen werkzaam zijn, door de leercapaciteiten van de leerlingen, de moeilijkheidsgraad van de leertaak en de onderwijsvaardigheden van de leerkrachten. Verschillen in leerprestaties tussen combinatie- en enkelvoudige klassen hebben we niet aangetroffen. Het gaat hierbij om de klassen 3 en 4 (groepen 5 en 6) (Veenman e.a., 1986; Veenman, Voeten & Lem, 1987; Veenman, Lem & Voeten 1987/88; Lem, Veenman & Voeten, 1989).

Uit de gegevens van deze onderzoeken en uit interviews met leerkrachten kan worden opgemaakt dat de meeste problemen in een combinatieklas kunnen worden teruggevoerd tot een vijftal gebieden: 1. het gebruik van de taakgerichte leertijd door de leerlingen, 2. het gebruik van de instructietijd door de leerkracht, 3. de klasse-organisatie en het klassemanagement, 4. het omgaan en het efficiënt inrichten van het zelfstandig werken, en 5. de school- of teamgerichtheid met betrekking tot combinatieklassen.

De onderzoeken die we tot dusver hebben uitgevoerd, kenmerken zich door hun beschrijvende en inventariserende aard. Er is onderzocht en beschreven hoe combinatieklassen functioneren in vergelijking met enkelvoudige klassen.

In een vervolgonderzoek is gepoogd de opgedane kennis en ervaringen te vertalen naar de concrete onderwijspraktijk. Door middel van een teamgerichte nascholingscursus 'Omgaan met Combinatieklassen' (OMC) is getracht de leerkrachten in combinatieklassen te informeren over en te bekwamen in effectieve en succesvolle onderwijsvaardigheden en hen te helpen deze in hun klas en school in te voeren om daarmee de kwaliteit van het onderwijs te verbeteren. Het onderzoek baseert zich op drie assumpties: a) onderzoeksgegevens kunnen als uitgangspunt genomen worden voor een kritische reflectie op het gegeven onderwijs en als zodanig dienen als een instrument voor schoolverbetering; b) onderzoeksgegevens kunnen voor de onderwijspraktijk van betekenis zijn als deze worden gezien als legitiem en als bruikbaar voor de praktijk; c) onderzoeksgegevens kunnen voor schoolleiders en leerkrachten dienstbaar zijn als zorgvuldige aandacht besteed wordt aan de presentatie, waarbij rekening gehouden wordt met de geaardheid van iedere school (Griffin & Barnes, 1986). In de volgende paragrafen wordt ingegaan op de onderzoeksvragen en de opzet, inhoud en uitvoering van de teamgerichte nascholingscursus 'Omgaan met Combinatieklassen'.

2 *Onderzoeksvragen*

Deze studie wil het uit onderzoek succesvol gebleken onderwijsgedrag met betrekking tot taakgerichte leertijd, effectieve instructie, efficiënt klassemanagement, zelfstandig werken en effectieve scholen vertalen naar de concrete onderwijspraktijk van de combinatieklas. Via een teamgerichte nascholingscursus wordt dit onderwijsgedragrovergedragen en geëvalueerd op implementatie. Door observaties, voorafgaande en na afloop van de cursus, wordt nagegaan of en in hoeverre het in de cursus geleerde een doorwerking heeft gevonden in de klas. Het onderzoek heeft de volgende vraagstelling: Leidt de gekozen vorm van teamgerichte nascholing tot implementatie van het geleerde in de schoolpraktijk van de combinatieklas (knowledge utilization)? Het effect van deze implementatie wordt in deze studie afgelezen aan de hand van een verbeterd taakgericht gedrag van leerlingen en onderwijsgedrag van leerkrachten volgens de in de nascholing gespecificeerde richtlijnen.

3 *De cursus omgaan met combinatieklassen*

3.1 *Samenwerkingsverband van de cursus*

De cursus is opgezet als een samenwerkingsverband tussen de Vakgroep Onderwijskunde, PABO, Schoolbegeleidingsdiensten en schoolteams[1]. Ieder van deze groepen participanten heeft een eigen deskundigheid en inbreng. De projectleden van de Vakgroep Onderwijskunde zijn vooral betrokken bij de inhoudelijke vormgeving en evaluatie van de cursus. De Pabo-docenten zijn vooral betrokken bij de cursusuitvoering. De schoolbegeleiders ondersteunen de cursisten op schoolniveau bij het invoeren van de kennis en vaardigheden die tijdens de cursus zijn opgedaan. De schoolteams brengen hun specifieke verwachtingen en behoeften in met betrekking tot de opzet, inhoud en begeleiding. Aldus is een (tijdelijke) werkorganisatie opgezet met netwerk-signatuur (Lem, Veenman & Nijssen, 1989).

De nascholingscursus is opgenomen in de regionale Wegwijzer voor nascholing en ondersteuning voor het schooljaar 1986-1987, aangeboden door de plaatselijke Pabo's en regionale schoolbegeleidingsdiensten. Met de 8 schoolteams die zich voor de nascholing hebben ingeschreven zijn voorafgaande aan de cursus twee intake-gesprekken gevoerd. In het eerste intake-gesprek is de cursus gepresenteerd. Er is ingegaan op de inhouden, de samenwerkingsrelaties, de klasseobservaties en de randvoorwaarden zoals tijdsinvestering, cursusplaats etc. Nadat de teams hun definitieve toezegging om aan de cursus deel te nemen, hebben gegeven, wordt enkele weken later in het tweede intake-gesprek ingegaan op specifieke behoeften van de schoolteams. Hierbij wordt ook aandacht besteed aan de rol van de schoolbegeleiders. In een eerder stadium is onderhandeld over de tijdsinvestering van de schoolbegeleiders en over de aard van de begeleiding. Deze onderhandelingen zijn gevoerd aan de hand van scenario's voor ondersteuning. Deze scenario's bestaan uit een aantal mogelijkheden om de ondersteuning van de leerkrachten, de teams en de cursusleiders zo veel mogelijk te garanderen en lopen van veel tot weinig begeleiding. Door gebrek aan begeleidingstijd wordt de ondersteuning door de vijf schoolbegeleiders als volgt ingevuld. Eén schoolbegeleider is regelmatig aanwezig bij de voorbesprekingen en bij de cursusbijeenkomsten, een andere (parttime) begeleider wordt door de bij de cursus betrokken Pabo-docent regelmatig op de hoogte gehouden. De andere begeleiders krijgen vóór elke cursusbijeenkomst een uitgebreide cursushandleiding zodat zij de tussentijdse ondersteuning kunnen voorbereiden. In die gevallen waarbij de schoolbegeleiders onvoldoende ondersteuning kunnen geven, kunnen de scholen een beroep doen op de leden van het onderzoeksteam. Een van de cursisten maakt van dit aanbod vrij intensief gebruik, anderen slechts incidenteel. De acht deelnemende schoolteams zijn verdeeld in vier cursusgroepen en vier cursusplaatsen (lokale scholen). Voor de concrete uitvoering van de cursus wordt verwezen naar Lem, Veenman, Nijssen & Roelofs (1988).

3.2 *De inhoud van de cursus*

In de cursus worden vijf thema's uitgewerkt: a. taakgerichte leertijd, b. doelmatige instructie, c. klassemanagement, d. zelfstandig werken, en e. schoolklimaat en schoolleiderschap. Deze thema's zijn gebaseerd op een aantal problematische aspecten van combinatieklassen dat in ons vorig onderzoek naar

voren kwam (Veenman e.a., 1986). De inhouden van de thema's zijn zoveel mogelijk ontleend aan empirische onderzoeksuitkomsten op het gebied van effectief onderwijzen en effectieve scholen.

a. *Taakgerichte leertijd.* Beschikbare tijd en taakgerichte leertijd zijn factoren die van grote invloed zijn op het onderwijsleerproces en uiteindelijk op het leren van de leerlingen (Walberg, 1986; Fraser, Walberg, Welch & Hattie, 1987). Door een grotere organisatorische complexiteit staat de leertijd in combinatieklassen in veel gevallen onder extra druk (Veenman, 1987). De leerkracht moet op verschillende momenten naar een andere (jaar)groep; het veelvuldig wisselen tussen (jaar)groepen kan de management- en overgangstijd tussen de activiteiten sterk verhogen; de leerlingen moeten in veel gevallen voor langere tijd zonder hulp en directe supervisie van de leerkracht kunnen werken; de heterogeniteit in de klas is extra groot, waardoor het afstemmen van de leerstof op het niveau van de leerlingen meer gecompliceerd is. De leertijd kan hierdoor minder doelmatig worden besteed. Tijdens de cursus wordt ingegaan op de taakgerichte leertijd, doelmatige tijdsbesteding, toegewezen tijd, succesniveau, afstemming van de leertaak. De leerkrachten worden aangemoedigd om strategieën toe te passen om de taakgerichtheid van de leerlingen te waarborgen of te verhogen. Bovendien zijn verschillende methoden gepresenteerd om de taakgerichtheid van de leerlingen te observeren.

b. *Effectieve instructie.* Leerlingen in combinatieklassen zijn, vergeleken met leerlingen in enkelvoudige klassen, vaak langdurig, zonder directe hulp en supervisie van de leerkracht bezig met individuele verwerking van leerstof. Bovendien blijkt dat tijdens individuele verwerking de taakgerichtheid van de leerlingen lager is dan tijdens activiteiten die door de leerkracht worden geleid (Veenman e.a., 1986). Wil de verwerkingsperiode voor de leerling succesvol zijn dan dienen aan de hieraan voorafgaande instructieperiode hoge eisen gesteld te worden. Recent onderzoek heeft een aantal bruikbare instructiemodellen opgeleverd (Rosenshine & Stevens, 1986; Good, Grouws & Ebmeier, 1983). De hoofdfasen van een effectief gebleken instructiemodel zijn: dagelijkse terugblik, ontwikkeling van nieuwe kennis, oefening onder nauwgezette begeleiding, zelfstandige verwerking, periodieke terugblik en veelvuldige terugkoppeling. In de cursus worden concrete toepassingen voor de combinatieklas beschreven. Met name wordt ingegaan op de vraag hoe de instructie zó kan worden gegeven dat de leerlingen zonder directe bemoeienis van de leerkracht verder kunnen waardoor de leerkracht zich kan concentreren op de andere groep. Videofragmenten illustreren het instructiemodel.

c. *Klassemanagement en organisatie.* Naarmate de klasseorganisatie gecompliceerder wordt, wordt het beroep dat wordt gedaan op de managementvaardigheden van de leerkracht groter. Een combinatieklas is een betrekkelijk gecompliceerde klasse-organisatie. De leerkracht moet allerlei managementvaardigheden in twee of meer groepen tegelijkertijd aanwenden. Klassemanagement wordt vanuit de cursus vanuit een viertal invalshoeken benaderd: 1. preventie van probleemgedrag, 2. didactische vaardigheden, 3. regels en afspraken en 4. de inrichting van de klas. Belangrijke managementvaardigheden die aan bod komen zijn: alert zijn, het spreiden van de aandacht, het erbij houden van de groep, het geven van verantwoordelijkheid aan de leerlingen, het zoveel mogelijk vermijden van lesonderbrekingen en het omgaan met regels en procedures (Kounin, 1970; Evertson, Emmer, Clements, Sanford & Worsham, 1984). Ook wordt aandacht besteed aan het doelmatig inrichten van het klaslokaal.

d. *Zelfstandig werken.* Leerlingen in combinatieklassen besteden naar verhouding veel tijd aan individuele verwerking van leerstof. Behalve dat hierdoor hoge eisen aan de instructie- en uitlegfase worden gesteld, betekent dit ook dat aan de inhoud en de organisatie van het zelfstandig werk veel aandacht moet worden besteed. De leerkrachten worden geconfronteerd met procedures die er toe bijdragen dat de taakgerichtheid van de leerlingen tijdens het zelfstandig werken wordt verhoogd, zoals bijvoorbeeld: a) de leerkracht besteedt meer aandacht aan de instructiefase en de fase van de geleide oefening, b) de leerkracht streeft er zoveel mogelijk naar om bij de leerlingen tijdens deze fasen 80% goede antwoorden of oplossingen te bereiken, c) de fase van het zelfstandig werken volgt onmiddellijk op de fase van de geleide oefening,

d) de leerstof waarmee de leerlingen tijdens het zelfstandig werken oefenen, sluit direct aan op de leerstof uit de instructie/uitlegfase,
e) de leerkracht begeleidt de leerlingen tijdens de eerste minuten van het zelfstandig werken om er zeker van te zijn dat alle leerlingen de leerstof begrepen hebben. Ook is aandacht besteed aan: de organisatie van het zelfstandig werk, dag- en weektaken, het omgaan met materialen, inleveren en corrigeren van gemaakte taken, en het organiseren van leerlinghulp. Voor de organisatie van het zelfstandig werken is onder andere gebruik gemaakt van het werk van Kierstead (1986).

e. *Schoolklimaat en schoolleiderschap.* Uit ons vorig onderzoek (Veenman e.a., 1986) blijkt dat op sommige scholen leerkrachten in combinatieklassen zich sterk geïsoleerd voelen en nauwelijks steun ondervinden van hun collega's in enkelvoudige klassen. Hierdoor worden overleg over specifieke onderwijskundige zaken, begrip voor elkaars situatie en continuering van beleid door de hele school bemoeilijkt. Combinatieklassen dienen de aandacht te hebben van het gehele team. Aandacht wordt besteed aan gegevens uit onderzoeken naar effectieve scholen, schoolmanagement, samenwerking van teamleden. Ook wordt de rol van de schoolleider benadrukt. Bij de invoering van combinatieklassen vervult de schoolleider een belangrijke rol (Veraart, 1988).

Bovenstaande thema's zijn geïntegreerd in een model gericht op de effectiviteit van de klas en de school. Dit bestaat uit de componenten: schoolleiderschap, schoolklimaat, leerkrachtgedrag, leerlinggedrag en leerresultaten (cf. Squires, Huitt & Segars, 1983). Voor verdere informatie over de inhoud van de cursus wordt verwezen naar Veenman, Lem en Nijssen (1988).

3.3 *Opzet en inrichting van de cursus*

De nascholingscursus omvat zeven bijeenkomsten van elk drie uur, die om de twee weken plaats vinden. De week tussen de bijeenkomsten is bestemd voor het uitvoeren van praktijkopdrachten. Voor het verbeteren van het onderwijsgedrag van leerkrachten in combinatieklassen is uitgegaan van een analyse van het gebeuren in de klas. Hierbij is gebruik gemaakt van de cyclus zoals toegepast in het project Achievement Directed Leadership (Research for Better Schools, 1985). Deze cyclus omvat de volgende fasen: 1. *Verzamelen van informatie.* Door middel van kwalitatieve en kwantitatieve observaties wordt informatie verzameld over leerkracht- en leerlinggedrag. De observaties richten zich voornamelijk op aspecten van leertijd, instructie, klassemanagement en zelfstandig werken. 2. *Vergelijken en analyseren.* Vervolgens wordt op basis van deze gegevens samen met de leerkrachten nagegaan of en in welke mate veranderingen gewenst zijn. 3. *Het opstellen van een veranderingsplan.* In overleg wordt een aantal veranderingsdoelen vastgesteld. Deze veranderingsdoelen kunnen per leerkracht verschillen. 4. *Uitvoeren van het veranderingsplan.* De gekozen veranderingsdoelen worden door de leerkrachten in hun eigen klassesituatie uitgeprobeerd en geëvalueerd. Hierbij worden ze ondersteund door collegaleerkrachten, schoolbegeleiders of projectmedewerkers. Hierna kan de cyclus opnieuw doorlopen worden.

De eerste twee fasen richten zich op het stellen van de diagnose en gaan aan de cursus vooraf. Vóór de cursus zijn bij de leerkrachten observaties verricht. Direct na de observaties zijn de gegevens teruggekoppeld. Na de behandeling van de eerste vier cursusthema's (leertijd, instructie, klassemanagement, zelfstandig werken) hebben de leerkrachten opnieuw terugkoppeling gekregen over de geobserveerde lessen, maar nu in het licht van de behandelde onderwerpen en met het oog op het opstellen van een veranderingsplan. Na de cursus zijn aan elke geobserveerde leerkracht de gegevens van zowel voor- als na-observatie in de vorm van cirkel- en staafdiagrammen gepresenteerd.

Bij de inrichting van de cursusbijeenkomst is gebruik gemaakt van de vijf trainingscompenenten die door Joyce en Showers (1980, 1988), Sparks (1983) en Stallings (1985) als effectief worden geïdentificeerd: 1. presentatie van theorie en achtergronden; 2. observerend leren en demonstratie; 3. oefening en toepassing in onbedreigde situaties; 4. gestructureerde terugkoppeling; en 5. coaching. De theorie en achtergronden zijn gepresenteerd in vijf cursuskaternen. Observerend leren en demonstratie gebeuren met behulp van videofragmenten, rollenspel en opdrachten uit de katernen. Oefening gebeurt door rollenspel en door het oefenen in de eigen klassesituatie. Terugkoppeling wordt gegeven door observa-

toren. Benadrukt is het belang van peerobservatie. In de cursus is coaching verzorgd door leden van het onderzoeksteam en op sommige scholen door de schoolbegeleider of schoolleider.

Het doel van de cursus OMC is niet om de leerkrachten te vertellen hoe ze moeten handelen, maar om ze instrumenten en concepten aan te reiken waarmee ze hun eigen onderwijsgedrag kunnen analyseren en verbeteren.

4 *Instrumentarium*

Om gegevens over het functioneren van de leerkrachten in combinatieklassen te verkrijgen wordt gebruik gemaakt van twee observatie-instrumenten. Het eerste observatie-instrument (COMMIT) wordt gebruikt om de tijdsbesteding van de leerlingen en leerkrachten te observeren en te registreren. Het tweede observatie-instrument (MIS) wordt gebruikt om vast te stellen of en in welke mate bepaalde management- en instructiegedragingen door de leerkracht worden aangewend.

COMMIT (Combinatieklassen Meetinstrument voor Instructie en Taakgericht gedrag) is een aanpassing en verfijning van het observatie-instrument BOLERO dat is gebruikt in het SVO-onderzoek 'Onderwijs in combinatieklassen' (Veenman e.a., 1986; Feenstra & Veenman, 1984). Het doel van COMMIT is om door middel van een specifieke vorm van 'time-sampling', namelijk de 'predominant activity sampling'-procedure (Tyler, 1979), een schatting te verkrijgen van de tijd die door de leerlingen en de leerkrachten in combinatieklassen aan verschillende activiteiten en settings wordt besteed. Deze procedure houdt in dat gedurende een relatief korte tijdseenheid wordt geobserveerd (7 seconden), waarna het gedrag dat gedurende deze tijdseenheid het meest dominant is, in de volgende 13 seconden wordt gecodeerd. De observator begint met het observeren en coderen van het gedrag van de eerste leerling en de leerkracht, vervolgens gaat hij over naar de tweede leerling en de leerkracht. In één minuut worden aldus drie verschillende leerlingen éénmaal en de leerkracht driemaal geobserveerd. Geobserveerd zijn lessen met een duur van 40 minuten.

Elke 20 seconden zijn vier aspecten van het klassegebeuren gecodeerd: 1. de taakgerichte leertijd van de leerlingen, 2. de groeperingsvorm waarin de leerlingen werken (setting), 3. de activiteiten van de leerkracht en 4. de groep waartoe de leerkracht zich richt. De observatiecategorieën zijn opgenomen in Tabel 1. Voor de analyse zijn sommige categorieën samengevoegd. De observaties zijn verricht met behulp van een schootcomputer (Epson HX-20). Voor dit doel is een speciaal programma ontwikkeld (Lassche, 1986).

Voorafgaande aan de observaties hebben de vier observatoren een training van circa 45 uur gevolgd. Bij de training is gebruik gemaakt van video-fragmenten en 'life'-observaties in combinatieklassen (3 reken- en 3 leestaallessen in 3 verschillende klassen). De interobservator-betrouwbaarheid, geschat met behulp van variantie-analyse (Winer, 1971), varieert van .82 tot 1.00 (met uitzondering van twee categorieën: procedurele activiteiten: .64, en geleide oefening: .72); mediaan: .98.

Het tweede observatie-instrument MIS (Management- en Instructie Schaal) is gebruikt om meer specifieke gegevens over de instructie- en managementvaardigheden van de leerkracht te verkrijgen. Het gaat hierbij niet om tijdschattingen, maar om schattingen of en in welke mate vooraf gedefinieerde leerkrachtgedragingen tijdens de les voorkomen. De MIS bestaat uit 30 items die de vorm hebben van bipolaire 5-puntsschalen. Bij de samenstelling van de schaal is gebruik gemaakt van ons eerder onderzoek en het werk van Evertson, Emmer, Sanford en Clements (1983) en Good e.a. (1983). De interobservator betrouwbaarheid (variantie-analyse) varieert van .67 tot .96 (mediaan: .86).

Op inhoudelijke gronden zijn de 30 items in 5 subschalen ingedeeld: 1. instructievaardigheden, 2. organisatie van de instructie, 3. gebruik van ruimte en materialen, 4. de afstemming van de instructie op de leerlingen, en 5. het omgaan met storingen. De resultaten van principale componentenanalyses op de items van de subschalen wijzen uit dat elke subschaal met behulp van één factor is te beschrijven. De (Cronbachs) alpha-coëfficiënten variëren van .66 tot .91.

Tot slot hebben de aan de cursus deelnemende leerkrachten een vragenlijst ingevuld over de opzet, inhoud en uitvoering van de cursus.

5 *Dataverzameling en -verwerking*

Voor de cursus hebben zich 8 schoolteams ingeschrevenr(7 voltallige teams en 1 subteam). Deze zijn afkomstig uit de regio's Boxmeer en Nijmegen. In totaal hebben 41 leerkrachten aan de cursus deelgenomen (periode januari-april 1987). Om praktische redenen (menskracht) zijn bij 17 leerkrachten observaties uitgevoerd[2]. Uit elk van de 8 teams zijn 2 à 3 leraren geobserveerd. Zij hebben lesgegeven in de combinaties 2/3 tot en met 7/8. Met inachtneming van spreiding over scholen en groepen zijn deze leerkrachten aselect gekozen. Deze leerkrachten vormen de experimentele groep (onderwijservaring 7-25 jaar). Om het effect van de cursus beter te kunnen vaststellen is eveneens een controlegroep samengesteld. Deze groep bestaat uit 9 leerkrachten in combinatieklassen, afkomstig uit 6 vergelijkbare scholen (onderwijservaring 5-22 jaar). Zij hebben niet aan de cursus deelgenomen en zijn niet op de hoogte van de inhoud ervan. Het onderzoek is aldus opgezet als een 'pretest-posttest control group design'. Voorafgaande aan de cursus zijn bij alle leerkrachten 2 reken- en 2 leestaallessen geobserveerd (november-december, 1986). Na afloop van de cursus zijn opnieuw 2 reken- en 2 leestaallessen geobserveerd (mei-juni, 1987). De observatieprocedures zijn bij beide groepen leerkrachten dezelfde geweest. De observaties hebben alle in de ochtend plaats gevonden. De observatiegegevens bestaan uit 136 observaties in de experimentele groep (17 klassen * 4 lessen * 2 perioden: voor- en nameting) en uit 72 observaties in de controlegroep (9 * 4 * 2).

Aangezien het bestaande lesrooster is aangehouden, is het niet altijd mogelijk geweest om de reken- en leestaallessen te scheiden. Zo komt het regelmatig voor dat een (jaar)groep met rekenen bezig is en de andere groep met taal, of dat binnen een (jaar)groep een aantal leerlingen met andere activiteiten bezig is dan andere leerlingen. Deze vorm van werken is inherent aan veel combinatieklassen. Dit betekent dat bij de analyses geen onderscheid is gemaakt tussen deze twee lessen, m.a.w. een analyse op het niveau van reken- of leestaalles heeft niet plaatsgevonden. De analyses zijn uitgevoerd op klas-, groeps- en leerkrachtniveau. Van een analyse op teamniveau is afgezien vanwege het geringe aantal geobserveerde leerkrachten per team. Een dergelijke analyse heeft wel plaatsgevonden inzake de invloed van netwerkvorming op het effect van de cursus (Lem, Veenman & Nijssen, 1989). De analyses zijn uitgevoerd met behulp van SPSS-X (Nie et al., 1986).

De observatiegegevens die zijn verkregen met behulp van het instrument COMMIT hebben per meting steeds betrekking op de gemiddelde tijdsbesteding van de hele klas, berekend over de vier geobserveerde lessen. Per observatiecategorie resulteert een tijdschatting die is uitgedrukt in minuten. De observatiegegevens van de MIS zijn omgezet in somscores (berekend door de itemscores binnen elk van de 5 subschalen op te tellen). De berekende item- en schaalscores hebben evenals de COMMIT-tijdschattingen betrekking op de gemiddelden van vier lessen.

Verschillen tussen de voor- en de nameting bij de experimentele groep zijn eenzijdig getoetst met behulp van een t-toets voor gepaarde waarnemingen. Verschillen tussen de experimentele groep en de controlegroep zijn getoetst met behulp van een t-toets voor verschilscores (nameting minus voormeting). Overwogen is om een covariantieanalyse uit te voeren. Hiervan is afgezien vanwege het vrijwillige karakter van de experimentele groep (zelfselectie kan correleren met scores op de voormeting), vanwege het gevonden significante verschil in taakgerichte leertijd tussen de twee groepen op de voormeting, en vanwege het relatief kleine aantal klassen/leerkrachten. Bij de toetsing is een significantieniveau van 5% aangehouden (eenzijdig).

Voor een uitgebreide beschrijving van de opzet, het instrumentarium en de dataverzameling wordt verwezen naar Lem e.a. (1988).

6 *Resultaten*

In Tabel 1 wordt een overzicht gegeven van de percentages tijdsbesteding in de verschillende categorieën van het observatie-instrument COMMIT.

Wanneer we de experimentele groep vergelijken met de controlegroep vóórdat de cursus start, vinden we een significant verschil met betrekking tot taakgerichte leertijd. In de combinatieklassen van de leerkrachten uit de controlegroep komen de leerlingen tot een taakgerichtheid van 74,3%, terwijl in de klas-

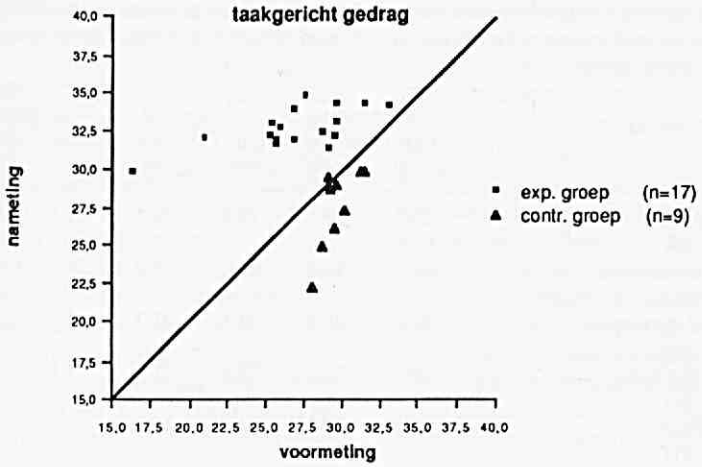
Tabel 1 *Overzicht van de gemiddelde tijdsbesteding (in percentages) van de experimentele en controlegroep op voor- en nameting en de toetsing (t-toets) van cursuseffecten met betrekking tot de tijdsbesteding van leerkrachten en leerlingen door middel van verschilscores (lesduur = 40 minuten = 100%)*

Observatie-categorieën	Voormeting Exp.	Contr.	Nameting Exp.	Contr.	Verschilscores Exp.	Contr.	*p*
A. LEERLINGGEDRAG							
Taakgericht	67.3	74.3	82.0	68.8	14.7	−5.5	**
Niet-taakgericht:							
– procedurele activiteiten	11.3	10.0	8.5	10.0	−2.8	−0.0	*
– wacht op leerkracht	3.8	3.3	1.0	3.5	−2.8	0.2	*
– niet met opdracht bezig	17.8	12.5	8.8	17.8	−9.0	5.3	**
B. SETTING							
Hoogste groep:							
– instructie	28.0	34.5	27.8	27.8	−0.2	−6.7	N.G
– verwerking	72.0	65.5	72.2	72.2	0.2	6.8	N.G
Laagste groep:							
– instructie	25.8	28.5	29.3	31.3	3.5	2.8	N.G
– verwerking	74.3	71.5	70.8	68.8	−3.5	−2.7	N.G
C. LEERKRACHT GEDRAG							
Instructie:							
– terugblik	1.3	0.8	4.3	0.5	3.0	−0.3	**
– uitleg/presentatie	32.8	41.5	26.8	35.0	−6.0	−6.5	N.G
– geleide oefening	10.0	7.5	12.0	5.3	2.0	−2.2	N.S
Begeleiding verwerking:							
– individuele hulp	26.8	18.5	29.8	19.3	3.0	0.8	N.G
– gerichte controle	6.5	5.5	10.0	7.8	3.5	2.3	N.S
Organisatie:							
– organisatie en leswisseling	15.8	16.0	13.3	18.8	−2.5	2.8	*
– niet-actief lesgedrag	6.8	10.3	4.3	13.5	−2.5	3.2	**
D. DOELGROEP							
– hele klas	10.8	7.8	11.5	7.5	0.7	−0.3	N.G
– hoogste groep	41.5	43.0	40.0	37.5	−1.5	−5.5	N.G
– laagste groep	41.0	39.0	44.8	41.5	3.8	2.5	N.G
– geen interactie	6.7	10.3	3.7	13.5	−3.0	3.2	*

Experimentele groep N = 17; Controlegroep N = 9. In de experimentele groep werd in twee klassen les gegeven aan één groep. Bij hercodering achteraf van de leerlingen in de hoogste en laagste groep traden ontbrekende scores op bij setting en doelgroep. Derhalve voor deze analyse: Setting en Doelgroep N = 15.
* $p<.05$; ** $p<.01$. N.S. = niet significant; N.G. = niet getoetst op implementatie.

sen van de leerkrachten uit de experimentele groep een taakgerichtheid van 67,3% wordt bereikt.

Uit de gegevens van Tabel 1 blijkt dat de OMC-cursus een opvallend effect heeft op het taakgericht gedrag van de leerlingen. De experimentele groep begint de cursus met een lager taakgericht gedrag van de leerlingen dan de controlegroep. Na de cursus blijkt de taakgerichte leertijd van de leerlingen uit de klassen van de experimentele groep, in vergelijking met de voormeting, significant te zijn gestegen. Het verschil tussen experimentele groep en controlegroep met betrekking tot de verandering in taakgericht gedrag (uitgedrukt in verschilscores) is significant ($p<.01$). Deze resultaten suggereren dat de leerkrachten uit de experimentele groep na de cursus beter in staat zijn om meer taakgericht gedrag bij de leerlingen te ontlokken. Uit Tabel 1 blijkt voorts dat de verhoging in taakgerichte leertijd gepaard is gegaan met een significante vermindering in wachttijd voor leerlingen ($p<.05$), met een vermindering van de tijd die



Figuur 1 *Gemiddelde taakgerichte leertijd per klas tijdens de voor- en nameting (in minuten)*

leerlingen besteden aan procedurele activiteiten ($p<.05$) en een vermindering van de tijd dat leerlingen niet met de opdracht bezig zijn ($p<.01$). Uit Figuur 1 (scatterplot) blijkt dat de spreiding in scores in de experimentele groep na de cursus sterk is teruggelopen. Alle leerkrachten uit de experimentele groep hebben de taakgerichte leertijd in hun klas kunnen verhogen.

In Tabel 1 wordt ook de gemiddelde tijd weergegeven die in de verschillende settings (instructie/uitleg en verwerking) door de twee groepen in de combinatieklas (de relatief hoogste en de relatief laagste groep) wordt besteed. In een instructiesetting heeft de leerkracht een leidende rol: hij legt uit, demonstreert, leidt het klassegesprek, presenteert, instrueert etc. De instructie is gericht tot de hele klas of tot een van beide groepen. In een verwerkingssetting heeft de leerkracht een begeleidende, een indirecte rol, de leerlingen werken zelfstandig of in groepjes. In combinatieklassen komt de verwerkingssetting het meest voor. Om na te gaan of in de twee settings sprake is van een verschillende 'holding power' op de aandacht van de leerlingen, zijn alle 7-seconden-observaties voor setting vergeleken met de scores voor taakgerichte leertijd.

In Tabel 2 wordt de gemiddelde taakgerichte leertijd per setting voor de experimentele en controlegroep weergegeven.

Na afloop van de cursus blijkt dat de leerlingen uit de experimentele groep in beide settings meer taakgericht zijn dan de leerlingen uit de controlegroep ($p<.01$). In de controlegroep is de taakgerichtheid in beide settings gedaald (voor de verwerkingssetting zelfs significant). Opvallend is dat in beide groepen, bij zowel voor- als nameting, de taakgericht-

Tabel 2 *Overzicht van de gemiddelde taakgerichte leertijd van de leerlingen (in percentages) per setting van de experimentele en controlegroep op voor- en nameting en de toetsing (t-toets) van cursuseffecten door middel van verschilscores (lesduur = 40 minuten = 100%)*

Settingcategorieën	Voormeting		Nameting		Verschilscores		
	Exp. groep	Controle groep	Exp. groep	Controle groep	Exp. groep	Controle groep	*p*
Instructie*Taakgericht	73.3	75.3	86.8	68.4	13.5	−6.9	**
Verwerking*Taakgericht	64.5	73.7	80.2	67.2	15.7	−6.5	**

Experimentele groep N = 17 (het oorspronkelijke aantal klassen geldt omdat voor deze analyse bij setting geen hercodering achteraf heeft plaatsgevonden in hoogste groep en laagste groep; vergl. Tabel 1); Controlegroep N = 9; ** $p<.01$

Tabel 3 *Overzicht van de gemiddelde somscores op de 5 subschalen van de Management & Instructie Schaal (MIS) van de experimentele en controlegroep op voor- en nameting en de toetsing (t-toets) van cursuseffecten door middel van verschilscores*

Subschalen	Voormeting		Nameting		Verschilscores		*p*
	Exp. groep	Controle groep	Exp. groep	Controle groep	Exp. groep	Controle groep	
Instructievaardigheden (7 items; alpha = .90)	23.1	24.1	28.2	23.3	5.1	−0.9	**
Organisatie van de instructie (5 items; alpha = .84)	15.7	16.6	20.0	17.4	4.3	0.8	**
Gebruik materiaal en ruimte (6 items; alpha = .66)	21.0	21.9	24.1	22.7	3.0	0.9	**
Afstemming van instructie (6 items; alpha = .86)	20.1	21.2	23.9	21.0	3.8	−0.2	**
Omgaan met storingen (6 items; alpha = .91)	18.3	20.1	23.6	19.7	5.3	−0.4	**

Experimentele groep N = 17; Controlegroep N = 9. Voor de 5-puntsschaal van elk afzonderlijk item geldt: 1 = meest negatieve score; 5 = meest positieve score; **$p<.01$.

heid tijdens de instructie hoger is dan tijdens de verwerking.

Tabel 1 geeft eveneens een vergelijking tussen de experimentele groep en de controlegroep voor wat betreft een aantal leerkrachtgedragingen. Het blijkt dat leerkrachten uit de experimentele groep significant meer tijd besteden aan een terugblik ($p<.01$). Belangrijker dan de gemiddelde hoeveelheid tijd die wordt besteed aan een terugblik, is het gegeven dat in de experimentele groep bij de voormeting 41% van de leerkrachten gebruik maakt van een terugblik, terwijl dit percentage op de nameting is gestegen tot 94%. In de controlegroep zijn deze percentages op voor- en nameting respectievelijk 44% en 33% (niet in tabelvorm opgenomen). Verder blijkt dat in de experimentele groep de tijd besteed aan wisselingen ($p<.05$) en aan niet actief lesgedrag ($p<.01$) significant is teruggelopen. Er zijn geen cursuseffecten te constateren met betrekking tot de tijd besteed aan 'geleide oefening' en 'gerichte controle'. Er zijn geen toetsingen uitgevoerd ten aanzien van de variabelen uitleg/presentatie, individuele hulp en aspecten van setting en doelgroep qua tijdsbesteding omdat verhoging ervan geen oogmerk van de cursus vormde. Kwalitatieve aspecten hieromtrent zijn gemeten met behulp van de MIS.

Tabel 3 geeft een overzicht van de scores op de subschalen van de MIS. Voor een totaaloverzicht van de afzonderlijke items en voor de wijze waarop de items over de 5 subschalen zijn verdeeld wordt verwezen naar Lem e.a. (1988).

Na vergelijking van de resultaten op de voor- en nameting, waarbij rekening wordt gehouden met de veranderingen in de controlegroep, kan worden vastgesteld dat de leerkrachten van de experimentele groep op alle aspecten van instructie- en managementgedrag vooruitgang hebben geboekt. Er is sprake van significant betere instructievaardigheden, een betere organisatie van de instructie, een adequater gebruik van materiaal en ruimte, en betere afstemming van de instructie op het niveau van de leerling en een meer doeltreffende omgang met storingen ($p<.01$). De items zijn ook afzonderlijk getoetst (Veenman, Lem & Roelofs, 1989). Uit deze toetsing blijkt dat de leerkrachten uit de experimentele groep op 24 van de 30 items een significante vooruitgang boeken (t-toets voor verschilscores); 18 items: ($p<.01$) en 6 items: ($p<.05$). De scores van de leerkrachten uit de controlegroep blijven nagenoeg gelijk.

7 *Discussie*

De vraag of deelname aan het schoolgerichte nascholingsprogramma 'Omgaan met Combinatieklassen' een positieve invloed heeft op de taakgerichte leertijd en het instructie- en managementgedrag van leerkrachten lijkt

door de gegevens van dit onderzoek, in ieder geval voorlopig, positief beantwoord te kunnen worden. In voortgezet onderzoek zal worden nagegaan of deze bevindingen herhaald kunnen worden als de cursus geheel verzorgd zal worden door PABO-docenten en schoolbegeleiders en of de cursus ook effecten heeft op leerlingniveau. Ook wordt overwogen effecten op langere termijn te meten.

Uit dit onderzoek blijkt dat de gemiddelde taakgerichte leertijd van de leerlingen uit de combinatieklassen van de experimentele groep ná de cursus significant hoger is dan vóór de cursus – respectievelijk 82% en 67%. Een verschil van 15% (of 5.9 minuten op een lesduur van 40 minuten). De taakgerichte leertijd in de controlegroep is op de voor- en nameting respectievelijk 74% en 69%. In deze groep loopt de taakgerichte leertijd iets terug, nl. 5% (of 2.2 minuten). Omdat op de voormeting een significant verschil is aangetroffen inzake de taakgerichte leertijd ten gunste van de controlegroep, moet bij deze uitkomsten worden aangetekend dat het mogelijk is dat de schoolteams die zich voor de cursus hebben aangemeld, behoefte hebben gehad aan een betere invulling en organisatie van de taakgerichte leertijd. Het verschijnsel van zelfselectie kan niet geheel worden uitgesloten.

Vergelijken we de gegevens uit dit onderzoek met die uit ons vorig onderzoek (Veenman e.a., 1986) dan blijken de verschillen in taakgerichte leertijd op de voormeting tussen beide groepen klassen klein: 71% in ons vorige onderzoek (12 combinatieklassen) en 69% in dit onderzoek (experimentele- én controlegroep: 26 combinatieklassen). In de groep enkelvoudige klassen (N = 12) uit het vorige onderzoek bedraagt de taakgerichte leertijd gemiddeld 77%.

Tussen klassen blijken zowel op de voor- als de nameting grote verschillen te bestaan. In de experimentele groep varieert de taakgerichte leertijd op de voormeting van 40% tot 83% en op de nameting van 75% tot 87%. In de controlegroep respectievelijk van 70% tot 79% en van 56% tot 75%. In de experimentele groep zijn de verschillen na de cursus kleiner.

Een belangrijk deel van het succes van de cursus moet worden toegeschreven aan de vormgeving en inhoud van de nascholing. De kern van de cursus wordt gevormd door vijf katernen die zijn gebaseerd op eerder onderzoek naar combinatieklassen en op onderzoek naar effectief onderwijsgedrag. Uit de gegevens die zijn verkregen uit de vragenlijsten blijkt dat de katernen door de leerkrachten zijn bestudeerd en gebruikt. Een groot gedeelte van de leerkrachten zegt dat de katernen en de cursusbijeenkomsten van veel nut zijn geweest omdat ze voorzien in concrete en praktische aanwijzingen. Ook de terugkoppelingsgesprekken op basis van de observaties zijn volgens de leerkrachten van veel belang geweest bij de invoering van het geleerde in de klassepraktijk. Voorts kan een deel van het succes van het nascholingsprogramma worden toegeschreven aan het feit dat geen van de in de cursus geoefende vaardigheden vreemd of nieuw zijn voor de leerkrachten. Waarschijnlijk bieden de encadrering en de achtergronden voor het gebruik van de vaardigheden een conceptueel kader van waaruit leerkrachten gerichte beslissingen kunnen nemen ten aanzien van hun onderwijzen.

Het onderzoek kent een aantal beperkingen. De mogelijkheid van zelfsectie is al genoemd. Effecten zijn gemeten direct na afloop van de cursus. Twee van de vier observatoren hebben geparticipeerd bij de cursusuitvoering en bij de terugkoppelingsgesprekken. Deze observatoren blijken beter in staat om de observatiegegevens en de cursusinhoud met elkaar in verband te brengen. Een nadeel is dat de verstrengeling van deze twee rollen de objectiviteit van de gegevens op de nameting kan hebben beïnvloed. Vergelijkingen tussen de vier observatoren met betrekking tot de gegevens op de nameting tonen echter geen significante verschillen of vertekening, maar de twijfel aan mogelijke beïnvloeding is hiermee niet geheel uitgesloten.

Ook de inhoud en uitvoering kent beperkingen. De toepasbaarheid van de cursus is als minder positief gewaardeerd door leerkrachten uit de kleutergroepen (de groepen 1 en 2). De inhoud is vooral afgestemd op de groepen 3 t/m 8. Ten aanzien van dit punt dient het programma bijgesteld te worden. Bij de uitvoering van de cursus zijn de twee Pabodocenten sterk ondersteund door projectmedewerkers. Deze projectmedewerkers hebben de inhoud en opzet van de cursus ontwikkeld en zijn daardoor en door hun ervaring met combinatieklassen uit eerder onderzoek beter op de hoogte van de praktijk van de combinatieklas en van onderzoeksgegevens over combinatieklassen, effectief onderwijzen en klas-

semanagement. Nascholingsprogramma's die worden geïnitieerd en ontwikkeld door een universiteit zijn volgens Wade (1984, 1985) over het algemeen effectiever dan programma's die door de scholen zelf of door de opleidingen worden geïnitieerd. Mogelijk komt dit door het gegeven dat professionals van buiten de school vaak meer tijd en mogelijkheden hebben om programma's te ontwikkelen, te toetsen en te presenteren. De volgende stap in het onderzoek zal dan ook gericht zijn op het trainen van Pabo-docenten en schoolbegeleiders met het doel dat zij het programma zelfstandig kunnen uitvoeren.

Hoewel in de cursus het belang en de mogelijkheden van coaching (het elkaar begeleiden door collega's, schoolleiders en schoolbegeleiders) is benadrukt, blijken hieraan een drietal problemen te kleven: praktische, psychologische en technische. Praktische problemen hebben betrekking op het feit dat de meeste van de aan de cursus deelnemende scholen klein zijn. Vijf van de 8 schoolteams hebben minder dan 6 leerkrachten. Deze leerkrachten en hun schoolleiders hebben een eigen klas. Ze zien nauwelijks mogelijkheden om zich vrij te maken om collega's in de les te bezoeken en te observeren. Psychologische problemen hebben betrekking op de weerstanden tegen observaties. Leraren zijn over het algemeen niet gewend dat collega's hun lessen observeren en bespreken. Problemen van technische aard hebben betrekking op de manier van elkaar begeleiden. Een aantal leerkrachten heeft de collega's slechts geobserveerd, terwijl andere leerkrachten uitgebreid zijn ingegaan op de observaties en samen hebben gezocht naar mogelijkheden voor verbetering. Bovendien blijken ook de schoolbegeleiders te weinig tijd te hebben de cursusbijeenkomsten en de klassen van de leerkrachten te bezoeken. Van de vijf bij de cursus betrokken schoolbegeleiders is slechts één begeleider in staat geweest de cursusbijeenkomst regelmatig te bezoeken en de teams op meer continue basis te coachen. De samenwerking tussen teams en schoolbegeleiders dient opnieuw overwogen te worden en de volle aandacht te krijgen. Ondanks deze mogelijke beperkingen kan worden gesteld dat een nascholingsprogramma zoals hier boven is beschreven succesvol kan zijn voor scholen met combinatieklassen.

Noten

1. Een samenwerkingsverband werd gesloten tussen de Facultaire Vakgroep Interdisciplinaire Onderwijskunde van de Katholieke Universiteit Nijmegen, de Katholieke Pedagogische Hogeschool Nijmegen-Arnhem, de Schooladviesdienst Nijmegen en Omgeving en de Onderwijsbegeleidingsdienst N.O. Brabant te Oss.
2. In de voormeting zijn bij 19 leerkrachten observaties verricht. Door observatievrees en ziekte zijn twee leerkrachten tijdens de nameting uitgevallen.

Literatuur

Evertson, C.M., E.T. Emmer, J.P. Sanford & B.S. Clements, Improving classroom management: An experiment in elementary school classrooms. *Elementary School Journal*, 1983, *84* (2), 173-188.

Evertson, C.M., E.T. Emmer, B.S. Clements, J.P. Sanford & M.E. Worsham, *Classroom management for elementary teachers*. Englewood Cliffs: Prentice Hall, 1984.

Feenstra, H. & S. Veenman, Het Bolero-systeem. *Computers op School*, 1984, *2*, 18-20.

Fraser, B.J., H.J. Walberg, W.W. Welch & J.A. Hattie, Syntheses of educational productivity research. *International Journal of Educational Research*, 1987, *11* (2), 147-252.

Good, T.L., D.A. Grouws & H. Ebmeier, *Active mathematics teaching*. New York: Longman, 1983.

Griffin, G.A. & S. Barnes, Using research findings to change school and classroom practices: Results of an experimental study. *American Educational Research Journal*, 1986, *23* (4), 572-586.

Inspectie Basisonderwijs, *Onderwijsverslag: Verslag van de staat van het onderwijs in Nederland*. 's-Gravenhage: Staatsuitgeverij, 1978.

Joyce, B.R. & B. Showers, Improving inservice training: The messages of research. *Educational Leadership*, 1980, *37* (6), 379-385.

Joyce, B.R. & B. Showers, *Student achievement through staff development*. New York: Longman, 1988.

Kierstead, J., How teachers manage individual and small-group work in active classrooms. *Educational Leadership*, 1986, *4* (2), 22-25.

Kounin, J.S., *Discipline and group management in classrooms*. New York: Holt, Rinehart & Winston, 1970.

Lassche, H., *Commit-programmatuur*. K.U. Nijmegen: Facultaire Vakgroep Interdisciplinaire Onderwijskunde, 1986.

Lee, J., Vertical grouping in the primary school. *School Organization*, 1984, *4*, 133-142.

Lem, P., S. Veenman & F. Nijssen, Netwerkvorming en teamgerichte nascholing. In: R. Vandenberghe & R. van der Vegt (Red.), *Onderwijsvernieuwing*. Lisse: Swets & Zeitlingen, 1989, 29-41.

Lem, P., S. Veenman & R. Voeten, Zeitnutzung und Schulleistungen in Mehrstufenklassen. *Zeitschrift für Pädagogische Psychologie*, 1989 (in druk).

Lem, P., S. Veenman, F. Nijssen & E. Roelofs, *Omgaan met combinatieklassen: Verslag van een teamgerichte nascholing*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs, Selecta-Reeks, 1988.

Nie, N. H. et al., Statistical Package for the Social Sciences - SPSS-X (Second Edition). New York: McGraw-Hill, 1986.

Research for Better Schools, *Achievement Directed Leadership: Training for teachers*. Philadelphia: Research for Better Schools, 1985.

Rosenshine, B. V. & R. Stevens, Teaching functions. In: M. C. Wittrock (Ed.) *Handbook of research on teaching*, 3rd. ed., 376-391. New York: MacMillan, 1986.

Santema, M., *Onderwijsbeleid en regio*. Leiden: Stafleu, 1974.

Sparks, G. M., Synthesis of research on staff development for effective teaching. *Educational Leadership*, 1983, *40*, 65-72.

Stallings, J., *A mastery approach to staff development*. Nashville, TN: Vanderbilt University, Peabody Center for Effective Teaching, 1985.

Squires, D. A., W. G. Huitt & J. K. Segars, *Effective schools and classrooms: A research-based perspective*. Alexandria: Association for Supervision and Curriculum Development, 1983.

Tyler, S., Time-sampling: A matter of convention. *Animal Behavior*, 1979, *27*, 801-810.

Veenman, S., Leertijd en basisvorming. In: J. van Kuyk (Red.), *Basisvorming in de basisschool*, 246-264. Tilburg: Zwijsen, 1987.

Veenman, S., P. Lem & F. Nijssen, *Omgaan met combinatieklassen: Een programma voor schoolverbetering*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs, Selecta Reeks, 1988.

Veenman, S., P. Lem & E. Roelofs, Training teachers in mixed-age classrooms: Effects of a staff development programme. *Educational Studies*, 1989, *15* (2), 165-180.

Veenman, S., P. Lem & M. Voeten, Time-on-task in mixed-age classes *Journal of Classroom Interaction*, 1987/88, *23* (2), 14-22.

Veenman, S., P. Lem & B. Winkelmolen, Active learning time in mixed-age classes *Educational Studies*, 1985, *11* (3), 171-180.

Veenman, S., M. Voeten & P. Lem, Classroom time and achievement in mixed-age classes *Educational Studies*, 1987, *13* (1), 75-89.

Veenman, S., P. Lem, M. Voeten, B. Winkelmolen & H. Lassche, *Onderwijs in combinatieklassen*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs, Selecta Reeks, 1986.

Veraart, C., *Schoolleiderschap en omgaan met combinatieklassen*. Doctoraalscriptie. Nijmegen: Facultaire Vakgroep Interdisciplinaire Onderwijskunde, 1988.

Wade, R. K., *What makes a difference in inservice teacher education: A meta-analysis of research*. Unpublished doctoral dissertation, University of Massachusetts, 1984.

Wade, R. K., What makes a difference in inservice teacher education: A meta-analysis of research. *Educational Leadership*, 1985, *41*, 48-54.

Walberg, H. J., Synthesis of research in teaching. In: M. C. Wittrock (Ed.) *Handbook of research in teaching*, 3rd. ed., 917-924. New York: MacMillan, 1986.

Winer, B. J., *Statistical principles in experimental design*. New York: McGraw-Hill, 1971.

Curriculum vitae

S. Veenman (1942) is universitair hoofddocent bij de Facultaire Vakgroep Interdisciplinaire Onderwijskunde van de Katholieke Universiteit Nijmegen. In februari 1975 promoveerde hij op het proefschrift 'Training op basis van interactieanalyse'. Zijn publicaties liggen op het terrein van goed en succesvol onderwijzen, schoolverbetering, opleiding en nascholing van leraren.

P. Lem (1953) studeerde onderwijskunde aan de K.U. Nijmegen. Hij was van 1982 tot 1988 als onderzoeker werkzaam bij de Vakgroep Onderwijskunde K.U. Nijmegen betreffende de SVO-projecten 'Onderwijs in combinatieklassen' en 'Omgaan met combinatieklassen'. Momenteel is hij als docent onderwijskunde verbonden aan de Pedagogisch Technische Hogeschool Nederland te Eindhoven.

E. Roelofs (1961) studeerde onderwijskunde aan de K.U. Nijmegen. Werkt sinds 1987 als onderzoeker binnen de Vakgroep Onderwijskunde K.U. Nijmegen aan het diepteproject 'Voortgezet technisch en begrijpend/studerend lezen' en aan de SVO-projecten 'Nascholing Leraren basisonderwijs' en 'Omgaan met combinatieklassen'. Is momenteel de dagelijkse projectleider van het SVO-project 'Omgaan met combinatieklassen (vervolgstudie)'.

Adres: Facultaire Vakgroep Interdisciplinaire Onderwijskunde, Katholieke Universiteit, Postbus 9103, 6500 HD Nijmegen.

Manuscript aanvaard 4-9-'89

Summary

Veenman, S., P. Lem & E. Roelofs, 'Dealing with mixed-age classrooms: Effects of a staff development programme.' *Pedagogische Studiën*, 1990, *67*, 45-58.

Increasingly, Dutch primary schools have no option but to introduce mixed-age classes because of the drop in pupil intake, and reduced staffing. A quasi-experimental, experimental-control group investigation was designed to test the effects of a staff development programme 'Dealing with Mixed-age Classes'. Research findings from mixed-age classes, effective teaching and classroom management and organization were translated into teacher behaviours. Following seven 3-hour workshops teachers in mixed-age classes implemented self-designed plans to increase selected research-derived teaching behaviours and pupils' time-on-task. Based on pre- and posttraining classroom observations, a significant treatment effect was found for pupils' time-on-task levels in mixed-age classrooms and for teacher behaviours regarding effective instruction, classroom organization and management.

De morfologische competentie van kinderen met leerproblemen

W. H. J. VAN BON, M. H. G. DIERX en J. J. M. P. KLERKX
Instituut voor Orthopedagogiek, Katholieke Universiteit Nijmegen

Samenvatting

Zeven- à achtjarige LOM-leerlingen hebben een geringere morfologische competentie dan even oude leerlingen van het Basisonderwijs. Deze achterstand betreft alle morfologische categorieën in gelijke mate. Er is geen verschil in type fouten. De achterstand strekt zich echter niet uit tot het hele domein van de cognitieve ontwikkeling: de LOM-leerlingen scoren weliswaar lager op de Coloured Progressive Matrices, maar het verschil in morfologische competentie blijkt daarmee niet verklaard te kunnen worden.

Andere factoren dan de intelligentie die bepaald wordt met de CPM en de ontwikkelingsachterstand die in de morfologietoetsen tot uiting komt, zullen moeten verklaren waarom de leerachterstand van de LOM-leerlingen groter is dan men op grond van deze beide zou voorspellen.

1 *Inleiding*

Vaak wordt van kinderen die leerproblemen[1] hebben gemeld, bijvoorbeeld door hun ouders of leerkrachten, dat zij achter lijken te zijn in taalontwikkeling (Dumont, 1974, 1982; Groenendaal, 1987; Hallahan & Kauffman, 1976). Dat is geen beoordelingsfout die, bijvoorbeeld, een halo-effect is van het leerprobleem. Uit groepsvergelijkend onderzoek met objectieve toetsingsprocedures blijkt immers dat kinderen met leerproblemen werkelijk een geringere taalvaardigheid hebben dan normaal lerende. Verschillen zijn gevonden in uiteenlopende aspecten van taalvaardigheid, van auditief-perceptuele prestaties (Pressman, Roche, Davey & Firestone, 1986) tot de weergave van verhalen (Ripich & Griffith, 1988) en de vaardigheid om het taalgebruik af te stemmen op de luisteraar (Spekman, 1981).

Bij onderzoek naar de Taaltests voor Kinderen (TvK) constateerde Van Bon (1984) dat er een verband is tussen de taalvaardigheid van een kind en het schooltype waar het zich bevindt, GLO of LOM. Discriminantanalyses lieten zien dat uit de scores die kinderen behalen op de taaltests tamelijk goed kan worden afgeleid welke schooltypen zij bezoeken. Het schooltype werd goed geïdentificeerd voor ongeveer 70% van de LOM-leerlingen en voor bijna 80% van de GLO-leerlingen, bij een gokkans van 50%. Van Bon concludeerde: "Taalvaardigheid kan dus een belangrijke determinant van schoolplaatsing zijn" (1984, p. 154).

Dat er een verband gevonden wordt tussen taalvaardigheid en leerproblemen is niet zo verwonderlijk. Er is een aantal redenen waarom taalcompetentie bepalend zou kunnen zijn voor iemands schoolloopbaan in het algemeen, die door het basisonderwijs in het bijzonder (Van Bon, 1984). Op de eerste plaats omdat het onderwijs grotendeels verbaal van aard is. De interactie tussen leerkracht en leerling gebeurt vooral door woorden. Zij stelt daarom niet alleen eisen aan het verbale aanpassingsvermogen van de leerkracht, maar ook aan de taalcompetentie van de leerling. Een tweede reden is gelegen in een belangrijk doel van het Basisonderwijs, leren lezen en schrijven, vaardigheden die een beroep doen op beheersing van de gesproken taal. Het is een gangbare opvatting dat lees- en schrijfproblemen het gevolg zijn van onvoldoende beheersing van de spreektaal op voor lezen en schrijven cruciale punten (Vellutino, 1979; Morais, Alegria & Content, 1987). Een derde argument is dat verbale codering (innerlijke of overte) een belangrijke faciliterende rol zou kunnen spelen bij het oplossen van problemen (Weisberg, 1980). Naarmate de taalvaardigheid geringer is, zal de oplossing van 'schoolproblemen' minder door verwoording ondersteund kunnen worden. Een vierde argument is dat de manier waarop een leerkracht de capaciteiten van een leerling inschat

en vermoedelijk ook zijn inzet ten behoeve van die leerling bepaalt, beïnvloed wordt door het taalgedrag van de leerling (Braun, 1976; Williams, 1970, 1976). De beduidende samenhang die men op grond van deze en andere argumenten zou verwachten tussen taalvaardigheid en schoolvorderingen, wordt ook gevonden. Van Bon (1984, 1986), bijvoorbeeld, vond dat gemiddeld 32% van de variantie in schoolprestaties wordt verklaard door variantie in de TvK-scores terwijl de samenhang met een 'niet-verbale' intelligentiemaat als Ravens Coloured Progressive Matrices slechts beperkt is (Van Bon, 1986). De bevinding dat leerproblemen samengaan met geringere taalvaardigheid sluit hierbij aan.

In onderzoek dat tot doel had taalachterstand van kinderen nader te specificeren werd de aandacht nogal eens gericht op hun vaardigheden op het terrein van de *morfologie*. Morfologie betekent 'vormleer' en wordt in de taalkunde gebruikt om het deel van de grammatica te benoemen dat handelt over vormverschillen tussen woorden, die samenhangen met verschillen van semantische en syntactische aard (Van Santen, 1984). Een vormverschil met een semantisch correlaat is GROEN – GROENIG. Een vormverschil met een vooral syntactische functie is (ik/hij) LOOP – LOOPT. Vormverschillen waarmee zowel semantische als syntactische verschillen corresponderen zijn HOOG – HOOGTE en SLAPEN – BESLAPEN. Woorden worden beschouwd als bestaand uit morfemen, de kleinste eenheden van een taal met een zelfstandige betekenis en/of grammaticale functie (Booij, Kerstens & Verkuyl, 1975). Booij et al. geven als voorbeeld HOEDEN dat bestaat uit de morfemen HOED ('hoofddeksel') en -EN ('meer dan één').

Het is gebruikelijk de morfologie te verdelen in *flexie* en *woordvorming*. Bij flexie of buiging wordt aan een woordstam (lexeem) een grammaticaal kenmerk toegevoegd, zoals bij verbuiging van zelfstandige naamwoorden (HOED-EN) en bij vervoeging van werkwoorden (ZOEK-T, ZOEK-EN). Bij woordvorming wordt in geval van derivatie of afleiding een woord gevormd uit een ander door er een affix aan toe te voegen (GROEN-IG, BE-SLAPEN) en in geval van compositie of samenstelling wordt een nieuw woord gevormd door twee woorden samen te voegen (BOSLOOP).

De afgrenzing van de gebieden van de morfologie ten opzichte van elkaar en ten opzichte van andere deelgebieden van de grammatica is verre van duidelijk en eenduidig (zie bijv. Van Santen, 1984). Zo rekent Extra (1978), zoals we verderop zullen zien, ook kennis van de lidwoorden die bij gegeven zelfstandige naamwoorden horen tot het domein van de morfologische kennis.

Hoewel het onderzoek naar de ontwikkeling van morfologische vaardigheden in het algemeen een ruimer gebied betrof (zie bijv. Schaerlaekens & Gillis, 1987), had dat bij kinderen met leerproblemen vooral betrekking op de beheersing van de flexie-regels en hun uitzonderingen.

De interesse voor de morfologische ontwikkeling van kinderen met leerproblemen wordt voor het Nederlandse taalgebied gelegitimeerd door enkele onderzoeksuitkomsten. Van der Leij (1983) vond dat van een aantal spreektaaltoetsen die voor morfologie (productie van verbogen vormen) opvalt door zijn hoge correlatie met leesscores en zijn vermogen om te differentiëren tussen goede en slechte lezers. Van der Wissel (1984) vond dat de Woordvormen-Productietest (WV-P) uit de TvK hoog correleerde met de dichotomie falen/niet-falen in het reguliere Basisonderwijs. Op grond van taalvaardigheidsbeoordeling verdeelde Pauw (1987) LOM-leerlingen in 'taalzwakke' en 'taalsterke' en vond een hogere WV-P-score bij de laatsten. Tenslotte laat factor-analyse van de TvK (Van Bon, 1984) zien dat de morfologie-tests een aparte status hebben: zij behoren tot de weinige tests die niet alleen gekarakteriseerd worden door een hoge samenhang met een algemene taalfactor, maar die ook een eigen factor definiëren.

Een vraag die bij de constatering van een ontwikkelingsachterstand van belang is, is of deze achterstand een beperkt gebied van kennis of vaardigheid betreft dan wel totaal is. Onder andere deze vraag stelden Wiig, Semel en Crouse (1973) toen zij kinderen met leerproblemen vergeleken met even oude normaal lerende. Zij gebruikten daarbij een aanpassing van een onderzoeksinstrument van Berko, de Wug-toets. Elk item van die toets bestaat uit twee nonsenseplaatjes. Bij elk plaatje zegt de proefleider een zin. De eerste zin bevat een pseudo-doelwoord (*Dit is een WUG*); de tweede zin is onaf en heeft zo'n

structuur (*Dit zijn twéé* . . . .) dat de proefpersoon wordt gestimuleerd een geflecteerde vorm van het doelwoord te produceren. De procedure kan louter inventariserend worden gebruikt; vaak echter wordt bepaald of de aanvullingen die kinderen geven overeenstemmen met de vormen die volwassen sprekers produceerden. Extra (1978) kenschetst deze procedure als een formele (in tegenstelling tot een informele zoals observatie zonder doelbewuste en gestructureerde ontlokking) en vanwege de pseudowoorden als een synthetische (in tegenstelling tot een naturalistische, waarin bestaande woorden worden gebruikt). Wiig et al. constateerden een kwantitatief verschil, lagere prestaties van de kinderen met leerproblemen, maar ook een kwalitatief verschil: de ordening van de morfologische categorieën naar moeilijkheid was voor de ene groep niet precies dezelfde als voor de andere. Wiig et al. concludeerden daarom dat er bij kinderen met leerproblemen geen sprake was van algehele achterstand, maar van individu tot individu variërende specifieke achterstand.

Morehead en Ingram (1973) en Johnston en Schery (1976) daarentegen vonden, niet met een dergelijke formele en synthetische procedure maar bij analyse van spontane uitingen, wel een tempoverschil, maar geen verschil in de volgorde waarin zich normaal ontwikkelende en in taalontwikkeling deviante kinderen de morfologische categorieën bemeesterden.

De vraag naar algeheel of specifiek verschil kwam ook aan de orde in het onderzoek van Extra (1978). Hij vergeleek met een formele procedure als die van Berko, maar naturalistisch door te vragen naar bestaande woorden, de morfologische vaardigheid van Turkse kinderen van 12 tot 15 jaar die het Nederlands als tweede taal leerden, met die van 4-jarigen voor wie het Nederlands de moedertaal was. Het instrument dat hij daarvoor samenstelde, de MopMov-test, had betrekking op 1. de lidwoordspecificatie bij zelfstandige naamwoorden en de daarmee samenhangende adjectiefvorming en keuze van voornaamwoorden (bijvoorbeeld *de* ballon, *dit* bed, grot*e* appel); 2. vorming van het meervoud en het diminutief (verkleinwoord) bij zelfstandige naamwoorden; 3. de vervoeging van werkwoorden. De rangcorrelatiecoëfficiënt tussen de beheersingsvolgordes van de onderscheiden grammaticale categorieën per groep was 0.85; volgens Extra een grote mate van overeenkomst, maar geen complete identiteit. Dat laatste baseert Extra ook op kleine percentuele verschillen. Die verschillen heeft hij echter niet getoetst en het lijkt niet onaannemelijk dat ze als toevalsfluctuaties moeten worden beschouwd. Hij concludeerde dat zich in beide taalleerprocessen vergelijkbare wetmatigheden voordoen (zoals de invloed van de formele complexiteit van een bepaalde regel en van de aanbodfrequentie), maar dat het leerproces in de tweede taal en dat in de moedertaal niet identiek zijn.

Over de vraag naar algehele of specifieke achterstand in morfologische kennis ging ook het onderzoek van Vogel (1983). Ze gebruikte een variant van Berko's taak (met pseudowoorden) om de kennis van morfologische regels te toetsen van kinderen met leerproblemen en normaal lerende. Ze vond een significant verschil ten nadele van de kinderen met leerproblemen. De overeenstemming tussen de moeilijkheidsrangordes van de onderscheiden morfologische categorieën was echter groot; de correlatie bedroeg 0.67. Vogel concludeerde dat de ontwikkeling later inzet en minder snel verloopt, maar kwalitatief niet anders is.

Het lijkt er dus op dat in geval van achterstand in morfologische ontwikkeling de beheersingsvolgorde misschien op punten afwijkt van die bij normale ontwikkeling, maar dat overeenkomst overheerst. Men zou dat dan ook verwachten bij Nederlandse kinderen met leerproblemen. Dit onderzoek is bedoeld om deze verwachting te toetsen door even oude Nederlandse kinderen met en zonder leerproblemen, gedefinieerd als leerlingen van het LOM-onderwijs en van het regulier Basisonderwijs, te vergelijken in hun morfologische kennis met behulp van twee bestaande tests, Extra's Mondeling Productieve Morfologische Vaardigheden Test (MopMov-test) en Van Bons Woordvormen-Productietest (WV-P). Beide tests maken gebruik van formele procedures, maar van naturalistische items. De WV-P heeft vooral betrekking op sterke en onregelmatige vormen; in de MopMov zijn ook de zwakke en de regelmatige vormen vertegenwoordigd. De items zijn te groeperen in afzonderlijke categorieën. Als kinderen met leerproblemen inderdaad minder morfologie-items goed beantwoorden is de vraag of dat

Tabel 1 *Enkele gegevens over de proefpersonen*

	BO-lln	LOM-lln
Leeftijd (mnd.)	98.1 (5.3)	98.4 (5.8)
CPM (percentiel)	58.4 (28.0)	33.9 (24.4)
Jongens/meisjes	30/10	28/12
Schoolprestaties:		
Een-Minuut-Test	41.5 (20.6)	18.5 (11.6)
Differentiële Zinnenleestest	178.6 (64.8)	119.3 (63.8)
Spellingtoets (prop. goed)	0.79 (0.18)	0.57 (0.22)
Schiedams Rekentoets (prop. goed)	0.78 (0.23)	0.48 (0.35)

geldt voor alle of slechts voor bepaalde categorieën. Als het laatste het geval is zou een karakterisering van specifieke achterstand licht kunnen werpen op de kern van het taalprobleem dat leerproblemen vergezelt en misschien veroorzaakt. Analyse zou kunnen gebeuren op grond van de formele complexiteit van de betreffende categorieën, de attentiewaarde van de morfologische elementen c.q. markeringen (bijvoorbeeld door hun positie in het woord, voor of achter) of de aanbodfrequentie (Extra, 1978).

2 *Methode*

2.1 *Proefpersonen*

Aan het onderzoek namen 80 kinderen deel, 40 leerlingen van het LOM-onderwijs, 40 van het Basisonderwijs (BO, groep 3 en 4).

Op de twee LOM-scholen hebben alle 7- en 8-jarigen van de twee laagste groepen meegedaan. Uit de leerlingen van de twee Basisscholen werden kinderen geselecteerd die naar geslacht, leeftijd en milieu het best overeenkwamen met de LOM-leerlingen. Tabel 1 geeft enkele karakteristieken van beide groepen proefpersonen.

2.2 *Onderzoeksinstrumenten*

Er werden twee morfologietests gebruikt, de WV-P en de MopMov-test.

1. De WV-P

De elicitatie-procedure maakt gebruik van een of meer plaatjes. De proefleider zegt daarbij een zin waarin het *doelwoord* is opgenomen in de woordenboekvorm of in een andere eenvoudige vorm (*Jan doet z'n best*). Vervolgens zegt hij een zin die de proefpersoon moet aanvullen met het doelwoord in een geflecteerde vorm die is bepaald door de structuur van de onvolledige zin (*Jan heeft z'n best....*). De test heeft 35 items met de volgende morfologische categorieën:

- Meervoud van zelfstandige naamwoorden (10 items)
- Diminutief van zelfstandige naamwoorden (3 items)
- Voltooid deelwoord (17 items)
- Comparatief en superlatief (5 items)

2. De MopMov-test

De procedure komt overeen met die van de WV-P. De test bestaat uit 3 hoofddelen die op hun beurt weer onderverdeeld zijn:

2.1. Lidwoordspecificatie bij zelfstandige naamwoorden en daarmee samenhangende pronominale specificatie en adjectiefmarkering:

- Lidwoordspecificatie (20 items, bijv. *Dit is een doos. Waar is m'n brief? M'n brief ligt op....* Criterium: DE / HET.)
- Pronominale specificatie (20 items, bijv. *Wat voor doos vind je het mooiste? Ik vind ....* Criterium: DEZE/DIT, DIE/DAT.)
- Adjectiefmarkering (20 items, bijv. *Zo'n doos is mooi. Dat is een ....* Criterium: MOOI/MOOIE.)

2.2. Meervouds- en diminutiefmarkering bij zelfstandige naamwoorden:

- Meervoud. De items zijn van het type: *Dit is een klok. Dat zijn er twee. Dat zijn twee ....*

Extra onderscheidt twee hoofdklassen van meervouden:

- Zonder modificatie van het lexeem of stamwoord (25 items, bijv. FLESSEN, SLEUTELS)
- Met lexeemmodificatie (20 items, bijv. VAAS-VAZEN, DAK-DAKEN)

- Diminutief. De items zijn van het type: *Dit is een emmer. Dat is ook een emmer. Maar dat is een heel klein ....*:
  - Zonder lexeemmodificatie (20 items, bijv. EMMERTJE, FLESJE)
  - Met lexeemmodificatie (5 items, bijv. BLAADJE, SCHEEPJE)

2.3. Verleden tijd- en voltooid deelwoordmarkering:

- Verleden tijd. De items zijn van het type: *De kinderen waren aan het zwemmen. Wat deden ze? Ik zag dat de kinderen ....*
- Voltooid deelwoord. De items zijn van het type: *Janneke is aan het zwemmen. Wat heeft zij gedaan? Zij heeft ....*

De verleden tijd en het voltooid deelwoord van dezelfde werkwoorden worden geëliciteerd:

- Zwakke werkwoorden (18 items). Bijv. BELLEN-BELDE, GEBELD; KNIPPEN-KNIPTE, GEKNIPT.
- Sterke werkwoorden (24 items). Bijv. GRAVEN-GROEF, GEGRAVEN; BREKEN-BRAK, GEBROKEN.
- Onregelmatige werkwoorden (8 items). Bijv. BRENGEN-BRACHT, GEBRACHT; ZOEKEN-ZOCHT, GEZOCHT.

Bij de analyse worden de scores op de overeenkomstige categorieën items van de WV-P en de MopMov-test bij elkaar opgeteld. De namen en afkortingen van deze 14 'morfologie-deeltoetsen' vindt men in Tabel 2.

Voor de detectie en statistische controle van eventuele intelligentieverschillen werd Ravens Coloured Progressive Matrices (CPM) afgenomen. Deze toets geeft in de normering van Van Bon (1986) een percentiele rang in verhouding tot de betreffende leeftijdsgroep.

Om beide groepen te karakteriseren in termen van schoolprestaties en voor de exploratie van de relatie tussen morfologie-(deel)tests en schoolvorderingen werden enkele toetsen afgenomen:

- Leestoetsen: De Een-Minuut-Test (EMT) (Brus & Voeten, 1973) en de Differentiële Zinnenleestest (DZT) (Dommerholt, 1981). Het aantal in 1 resp. 3 minuten goed gelezen woorden wordt bepaald.
- Spellingtoetsen: Een spellingtoets van eigen makelij, bestaande uit 15 woorden die op basis van hun foneemstructuur via eenvoudige foneem-grafeemcorrespondentieregels geschreven kunnen worden en 15 in lengte en (zoveel mogelijk) letters overeenkomstige woorden waarbij gekozen moet worden tussen homofone spellingen (ei/ij, au/ou, -g/-ch, f-/v-, s-/z-) en die dus schrijfwijze-kennis vragen voor competente spelling. De items werden gekozen uit de woordenlijsten van Van der Geest, Van der Kooij, Van Rossum, Ruijssenaars, Swüste en Westerbeek (1978), zodat de woorden geschikt mogen worden geacht voor deze proefpersonen. Het doelwoord werd eerst in een verduidelijkende zin gezegd en vervolgens afzonderlijk herhaald.
- Rekentoetsen: De testbladen 1 en 2 van de Schiedamse Rekentoets (Heesen, Strelitski & Van der Wissel, 1970) werden gebruikt. Deze bevatten optel- en aftreksommen tot en met 20.

2.3 *Procedure*

De individuele toetsafname (WV-P, MopMov, CPM, EMT & DZT) werd over twee dagen verdeeld, tweemaal drie kwartier. De spellingtoets en de rekentoets zijn door de eigen leerkrachten klassikaal afgenomen.

3 *Resultaten en discussie*

De gangbare opvatting over leerproblemen voorspelt geen verschil in niet-verbale intelligentie tussen kinderen met en zonder leerproblemen. Volgens de gebruikelijke definitie zijn leerproblemen immers leerprestaties die lager zijn dan men zou verwachten. De verwachting baseert men op een intelligentieschatting (en op de afwezigheid van bepaalde negatieve condities). Daarnaast hanteert men vaak het additioneel criterium dat de intelligentie gemiddeld moet zijn, althans niet lager dan een bepaalde waarde mag zijn. Als men de overtuiging is toegedaan dat verbale aspecten van de intelligentie veelal van beperkt niveau zijn, moet men zijn verwachting baseren op niet-verbale aspecten van de intelligentie. Nemen we de CPM als aanduiding voor de niet-verbale intelligentie, dan zouden we, in de mate waarin zo'n leerproblemenconcept aan LOM-plaatsing ten grondslag ligt, moeten vinden dat de LOM-leerlingen minstens even hoge CPM-scores hebben als BO-leerlingen.

Tabel 2 *Gemiddelde 'percentages goed' op de morfologie-deeltoetsen*

	BO-lln	LOM-lln
Meervoud zonder lexeemmodificatie (MvZLm)	0.89	0.78
Meervoud met lexeemmodificatie (MvMLm)	0.79	0.53
Diminutief zonder lexeemmodificatie (DZLm)	0.97	0.88
Diminutief met lexeemmodificatie (DMLm)	0.75	0.42
Verleden tijd, zwak (VtZw)	0.77	0.43
Verleden tijd, sterk (VtSt)	0.46	0.19
Verleden tijd, onregelmatig (VtOnr)	0.48	0.25
Voltooid deelwoord, zwak (VdZw)	0.98	0.91
Voltooid deelwoord, sterk (VdSt)	0.87	0.73
Voltooid deelwoord, onregelmatig (VdOnr)	0.91	0.79
Comparatief, Superlatief (Comp, Sup)	0.68	0.51
Lidwoordspecificatie (Lidw)	0.92	0.77
Pronominale specificatie (Pron)	0.87	0.69
Adjectief-specificatie (Adj)	0.91	0.75

De gemiddelde percentiele rangen van beide groepen op de CPM (zie Tabel 1) verschillen echter significant (t-toets, $p<0.01$) in het nadeel van de LOM-groep. Er is dan ook een significante correlatie (.46) tussen CPM-score en schooltype. De gemiddelde percentiele rang van de BO-leerlingen komt overeen met wat men zou verwachten bij een toevalssteekproef uit de normeringspopulatie. De oorzaak van het verschil is dus niet dat de BO-groep door toeval bestaat uit bovennormaal intelligente kinderen. Voor de centrale vraagstelling van dit onderzoek houdt deze bevinding in dat het een bron kan zijn van rivaliserende verklaringen: predicties op basis van verschil in morfologische competentie zouden kunnen worden verklaard door het verschil in CPM-scores. Daarom wordt waar nodig steeds de CPM-score als controle-variabele opgenomen.

De LOM-leerlingen scoren op alle morfologiedeeltoetsen lager dan de BO-leerlingen (Tabel 2). Een MANOVA met schooltype als onafhankelijke variabele en de scores op de 14 deeltoetsen als afhankelijke variabelen gaf een significante multivariate $F(14,65) = 3.47$ ($p<0.01$); ook alle univariate F's zijn significant. De groepen verschillen dus inderdaad in morfologische competentie. Dit onderscheid wordt niet verklaard door het verschil in intelligentie, want corrigeert men daarvoor door de CPM-percentielscore als covariaat in de toetsing op te nemen dan blijft het groepsverschil significant (multivariate $F(14,64) = 2.33$, $p = 0.01$); alle univariate F's zijn significant. Deze bevinding is overigens niet vreemd want de correlatie tussen de CPM en de morfologie-deeltoetsen is nogal laag, gemiddeld .30.

Tabel 3 *Discriminant-analyse met schooltype als criterium en de morfologie-deeltoetsen als predictoren (cursief CPM)*

	Voorspeld schooltype	
	BO	LOM
Werkelijk schooltype		
BO	77.5 *62.5*	22.5 *37.5*
LOM	20.0 *27.5*	80.0 *72.5*
	Correcte klassificatie: 78.8 *67.5*	

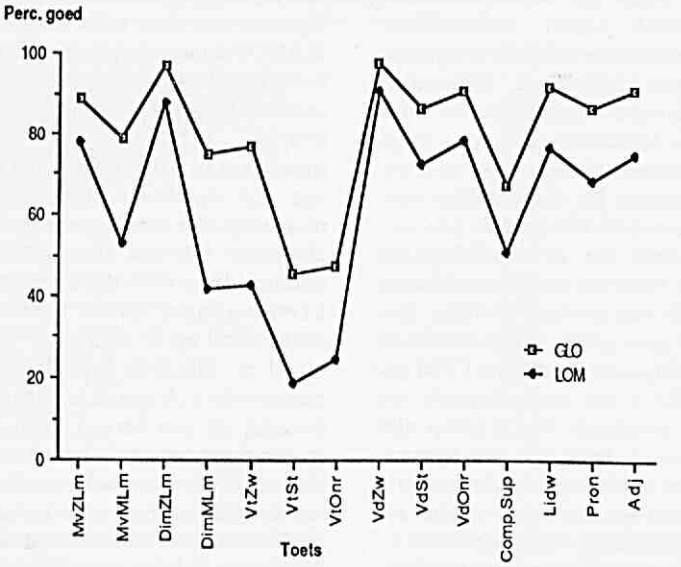
Het is dan ook tamelijk goed mogelijk uit de morfologie-deeltoetsscores af te leiden van welk schooltype een leerling afkomstig is. Van bijna 79% van de proefpersonen wordt het schooltype goed geïdentificeerd bij discriminant-analyse met de morfologie-deeltoetsen als predictoren (Tabel 3). Dat is aanzienlijk beter dan men op basis van toeval zou verwachten (50%), en ook beter dan met de CPM als predictor (68%). Voegt men de CPM als predictor toe aan de morfologie-deeltoetsen dan verbetert de predictie slechts in beperkte mate (tot 81%). De identificatie (of predictie, zo men wil) is echter niet volmaakt: rond 20% van elke groep heeft morfologiescores die kenmerkend zouden zijn voor de andere groep. Aan de geschatte percenta-

ges mag echter maar een beperkte waarde worden toegekend omdat discriminantanalyse ernaar tendeert de predictie te optimaliseren door kans-kapitalisatie. Bovendien zijn de scores waarop de predictie berust uiteraard enigermate onbetrouwbaar, wat enige overlap van de scoreverdelingen van de groepen kan veroorzaken. De enige stellige conclusie die men aan deze uitkomsten kan verbinden is dan ook dat de morfologische competentie van kinderen met leerproblemen geringer is dan die van normaal lerenden. Bovendien lijkt dat geen gevolg te zijn van de de niet-verbale intelligentie die met de CPM gemeten wordt. Het is dus waarschijnlijk een werkelijk 'talig' fenomeen. Het is echter niet uitgesloten dat er kinderen zijn met leerproblemen die zich in morfologische competentie niet onderscheiden van leerlingen van het reguliere basisonderwijs, en omgekeerd.

Nu geconstateerd is dat kinderen met leerproblemen inderdaad in de regel een achterstand hebben in morfologische competentie is de vraag gewettigd of die achterstand algemeen is of dat er sprake is van specifieke uitval. Algemene achterstand zou inhouden dat de relatieve moeilijkheid van de onderscheiden deelgebieden voor beide groepen dezelfde is, specifieke uitval dat sommige deelgebieden de LOM-leerlingen naar verhouding veel problemen geven. Er is echter geen specifieke uitval. De correlatie tussen BO en LOM in de moeilijkheid van de deeltoetsen – d.i. de correlatie tussen de kolommen van Tabel 2 – is zeer hoog, 0.96 (zowel Pearsons Product-Moment- als Spearmans rangcorrelatie coëfficiënt) en Figuur 1 illustreert dat: de moeilijkheidspatronen komen sterk met elkaar overeen. Er is geen aanwijzing dat bepaalde deeltoetsen de LOM-leerlingen uitzonderlijke moeite geven. Dat blijkt ook uit de volgende berekening. Een factoranalyse over de morfologie-deeltoetsen (gepoolde binnengroepscorrelatiematrix, hoofdassenmethode) geeft twee factoren met een eigenwaarde groter dan 1. De eerste ongeroteerde factor is te beschouwen als de algemene factor die laat zien hoeveel alle toetsen gemeenschappelijk hebben. Hij verklaart veel variantie, 58%, in de toetsen. Dus is er veel gemeenschappelijks, door taak-interne factoren, door externe factoren, of door beide. De toets die de algemene factor het best vertegenwoordigt is MvMLm, die op die factor een lading heeft van .84, en een onbeduidende lading op de tweede factor. Opname van deze toets als covariaat in een MANOVA met schooltypen als onafhankelijke variabele en de overige morfologiedeeltoetsen als afhankelijke variabelen, geeft een multivariate $F(14,65) = 1.24$. Deze is niet significant ($p = 0.27$). Ook alle univariate F's zijn niet significant. Het groepsverschil in morfologische competentie verdwijnt dus na eliminatie van wat aan de toetsen gemeenschappelijk is. Ook hieruit volgt dat er bij de LOM-leerlingen sprake is van een algehele achterstand en er daarnaast geen specifieke uitval is. Mochten bepaalde morfologische categorieën zich van de andere onderscheiden doordat zij een beroep doen op specifieke (psycholinguïstische) bekwaamheden, dan differentiëren deze bekwaamheden niet tussen de beide groepen proefpersonen.

Evenmin is er zo'n verschil tussen deze beide groepen en de twee groepen proefpersonen die Extra (1978) onderzocht. De correlatie tussen de deeltoetsmoeilijkheidsgraden die Extra publiceerde voor de 4-jarigen die het Nederlands als moedertaal leerden en die uit dit onderzoek is 0.92 voor de BO-leerlingen en 0.88 voor de LOM-leerlingen. De correlatie met de moeilijkheidsgraden voor de Turkse kinderen die Nederlands leerden is 0.86 voor de Basisschoolleerlingen en 0.83 voor de LOM-leerlingen. Het bekwaamheidsverschil tussen de groepen betreft alle morfologische categorieën dus in overeenkomstige mate. De conclusie lijkt gewettigd dat de ontwikkeling van de morfologische competentie over de gehele linie, d.w.z. op de onderscheiden deelgebieden tegelijk plaatsvindt en dat geen deelgebied c.q. morfologische categorie bij een groep in ontwikkeling relatief achterblijft of vooruit is.

Een van de manieren waarop een achterstand zich zou kunnen uiten is door een tendens tot overgeneralisatie of overregularisatie, wat inhoudt dat een regel ten onrechte wordt toegepast. Er is dan geen verschil in de regelmatige categorieën, wel in de onregelmatige die als regelmatige behandeld worden. Voorbeelden zijn SLIEP → SLAAPTE en EIEREN → EIEN. Dit verschijnsel is vooral bekend uit de literatuur over de ontwikkeling in de kennis van het sterke werkwoord. Volgens Schaerlaekens en Gillis (1987; zie ook Dijkstra & Kempen, 1984) gebruikt het kind aanvankelijk vooral



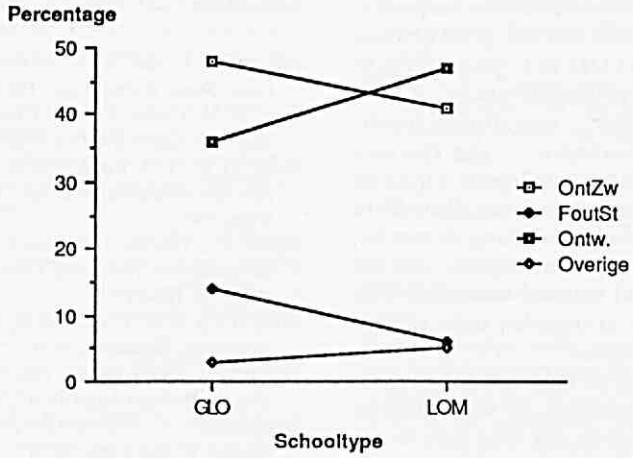
Figuur 1 *Gemiddelde 'percentages goed' op de morfologie-deeltoetsen*

de juiste vormen van de sterke werkwoorden: KWAM, AT, LIEP. Deze werkwoorden benoemen vaak voorkomende handelingen en gebeurtenissen en zijn hoogfrequent in het taalaanbod waaraan kinderen blootstaan. Hun vormregelmaat is dermate gering dat er geen regels uit kunnen worden afgeleid. Gaandeweg is het kind echter zo vaak geconfronteerd met zwakke werkwoorden dat het er regels voor heeft kunnen ontwikkelen. Deze gaat het ook toepassen in plaats van de tot dan correcte, vermoedelijk als afzonderlijke lexicale items geleerde vormen van sterke werkwoorden: KOMDE, EETTE, LOOPTE, GEKOMD, GEEET, GELOOPT. In dit stadium treedt niet de vocaalwisseling op die kenmerkend is voor het sterke werkwoord. Dat gebeurt wel in het volgende stadium: KWAMDE, ATTE, LIEPTE, GEKWAMD, GEAT, GELIEPT. In het laatste stadium worden weer de correcte sterke vormen gebruikt.

Hierboven was al geconstateerd dat er alleen een algemeen niveau-verschil is tussen de beide groepen leerlingen en geen verschil op specifieke morfologische categorieën, dus ook geen verschil m.b.t. regelmatige resp. onregelmatige en sterke werkwoorden. Omdat het echter niet uitgesloten is dat de fouten van LOM-leerlingen van een andere soort zijn dan die van de BO-leerlingen, werden de responsies op de werkwoords-items van de MopMov geanalyseerd. De fouten werden geklassificeerd als Onterecht Sterk (SNOEPTE → SNOOP), Foutief Sterk (HING → HONG), Onterecht Zwak (HING → HANGDE), Foutief Zwak (KAMDE → KAMTE), Zwak & Sterk (VLOOG → VLOOGDE) en Ontwijkend (*De kinderen waren aan het zwemmen. Wat deden ze? Ik zag dat de kinderen ...* AAN HET ZWEMMEN WAREN). Figuur 2 laat de procentuele verdeling van de werkwoordsfouten over deze categorieën zien; de minder vaak voorkomende categorieën zijn samengenomen.

Corrigeren we voor het totale aantal fouten tegen de werkwoorden, dan blijkt er geen verschil te zijn tussen de leerlinggroepen in het aantal fouten van de verschillende categorieën (MANOVA, $F(5,73) = 1.67$, $p = 0.15$). We moeten dus constateren dat de LOM-leerlingen meer fouten maken dan de BO-leerlingen, maar dat op dezelfde manier doen.

De grenzen van de taalachterstand doorsnijden het onderzochte morfologische domein kennelijk niet, maar liggen daarbuiten. Is de achterstand beperkt tot morfologische kennis? De uitkomsten van onderzoek dat we in de inleiding noemden (Pressman et al., 1986; Ripich & Griffith, 1988; Spekman, 1981) suggereren dat het gebied van achter-



Figuur 2 *Werkwoorden MopMov, foutentypen*

stand ruimer is. Ook Van der Wissel (1984), die de doelgroep enigszins afwijkend definieerde en selecteerde als 'falende GLOleerlingen', vond een verschil t.o.v. niet-falende in meerdere aspecten van taalvaardigheid, maar niet in alle. De samenhang van de Woordvormen-Productietest en de Zinsbouw-Productietest met de dichotomie falen vs. niet-falen was hoger dan die van de Test voor Verzwegen Betekenis. Van der Wissel concludeerde uit deze en andere uitkomsten dat de taalachterstand van falende GLO-leerlingen aldus gekarakteriseerd zou kunnen worden: "Opgaven, verbaal aangeboden en verbaal te beantwoorden, handelende over factische informatie die alleen berust op louter verbaal te verkrijgen informatie, met name ook informatie over taalvormaspecten, vormen het grote struikelblok voor de falers" (1984, blz. 215). Onze bevindingen passen in Van der Wissels karakterisering, maar kunnen hem niet op alle fronten verifiëren door de beperkte omvang van het onderzochte taaldomein.

De ontwikkelingsachterstand betreft echter niet het gehele cognitieve domein, want door correctie voor de CPM-percentielscores wordt het morfologie-verschil niet geëlimineerd. Dientengevolge moesten we (eerder al) concluderen tot een waarschijnlijk werkelijk 'talig' fenomeen. Was het verschil in CPM-scores echter in overeenstemming geweest met dat in morfologie-scores, dan was aan een essentiële voorwaarde voor het gebruik van de aanduiding 'leerproblemen' (normale intelligentie) nog minder voldaan dan nu al het geval was. Stanovich, Nathan en Vala-Rossi (1986) wijzen er overigens op dat de manier waarop dyslexie – maar de uitspraak is generaliseerbaar naar leerproblemen – gedefinieerd is ('normale intelligentie'), een ontwikkelingsachterstand die zich uitstrekt tot met de intelligentie gecorreleerde taken bij voorbaat uitsluit.

Zijn de leerprestaties wel in overeenstemming met de achterstand in morfologische competentie? Tabel 1 laat zien dat de schoolprestaties van de BO-leerlingen, zoals te verwachten is, beter zijn dan die van de LOM-leerlingen. Dit verschil is significant (multivariate $F(4,74) = 11.86$, $p<0.01$; alle univariate F's zijn significant) en verdwijnt niet bij correctie voor de algemene factor in de morfologie-deeltoetsen (multivariate $F(4,73) = 4.23$; het verschil in DZT-scores is niet meer significant). De achterstand die in de morfologie-toetsen tot uitdrukking komt is dus niet de enige bepalende factor voor de leerachterstand. Als het de uiteindelijke oorzaak is, moeten andere factoren de leerachterstand vergroot hebben. Naar die andere factoren kan men slechts gissen.

Het verschil in leerprestaties wordt evenmin in zijn geheel verklaard door de geringere niet-verbale intelligentie van de LOM-leerlingen, want met de CPM als covariaat blijft het verschil significant (multivariate $F(4,73) = 7.38$, $p<0.01$; alle univariate F's zijn significant). Dus ook als men statistisch corrigeert voor het intelligentieverschil tussen

beide groepen blijft het verschil in leerprestaties bestaan. Was dat niet het geval geweest dan was het gehele LOM-BO-verschil terug te brengen tot een intelligentieverschil. Kinderen met leerproblemen – operationeel gedefinieerd als LOM-leerlingen – zijn dus niet eenvoudigweg minder intelligente kinderen met alle verdere kenmerken van dien. Deze conclusie boet echter in aan belang als men bedenkt dat de niet-verbale intelligentie, ook die welke met de CPM bepaald wordt (zie Van Bon, 1986), maar in beperkte mate samenhangt met leerprestaties.

Een probleem bij dit type studies betreft de bepaling van de causaliteit. Er wordt immers slechts een samengaan van twee deficits geconstateerd. Het is niet uit te sluiten dat de achterstand in morfologische ontwikkeling een gevolg is van de leerproblemen en de daarmee samenhangende LOM-plaatsing. Leerproblemen bestaan immers meestal deels of geheel uit lees- en spellingproblemen die ook repercussies hebben voor de mondelinge taalcompetentie. Bovendien is te verwachten dat ook de LOM-plaatsing een negatief effect heeft op de taalontwikkeling; het verblijf in een lagere klas is immers al nadelig voor de taalvaardigheid (Van Bon, 1984). Een mogelijkheid om de onzekerheid over de causaliteitsrichting te beperken is groepen kinderen te selecteren met dezelfde schoolprestaties maar van verschillende leeftijd (zie bijv. Stanovich e.a., 1986). Superioriteit van de jongere groep in taalvaardigheid ten gevolge van schoolervaring (meer lees- en schrijfervaring) en schoolplaatsing ligt dan minder voor de hand.

Noot

1. Met leerproblemen worden hier die tegenvallende schoolprestaties bedoeld waarvoor ook wel de termen 'leerstoornissen' en 'primaire leerproblemen' worden gebruikt.

Literatuur

Bon, W. H. J. van, *Aspecten van taalvaardigheid*. Lisse: Swets & Zeitlinger, 1984.

Bon, W. H. J. van, *Raven's Coloured Progressive Matrices*. Lisse: Swets & Zeitlinger, 1986.

Booij, G. E., J. G. Kerstens & H. J. Verkuyl, *Lexicon van de taalwetenschap*. Utrecht: Het Spectrum, 1975.

Braun, C., Teacher expectation: sociopsychological dynamics. *Review of Educational Research*, 1976, *46*, 185-213.

Brus, B. Th. & M. J. M. Voeten, *Een-Minuut-Test*. Nijmegen: Berkhout, 1973.

Dijkstra, A. & G. Kempen, *Taal in uitvoering*. Groningen: Wolters-Noordhoff, 1984.

Dommerholt, J., *Differentiële Zinnenleestest*. Groningen: Wolters-Noordhoff, 1967.

Dumont, J. J., *Leerstoornissen. Deel 3: controversen en perspectieven*. Rotterdam: Lemniscaat, 1982.

Dumont, J. J., *Lees- en spellingsproblemen*. Rotterdam: Lemniscaat, 1984.

Dumont, J. J., Taalontwikkelingsachterstanden en taalstoornissen als voorwaarde voor het ontstaan van leerstoornissen. In: *Preventie van spraakstoornissen en taalachterstand*. Nijmegen: NVLF, 1974.

Extra, G., *Nederlands van buitenlanders*. Nijmegen: Academisch Proefschrift, 1976.

Geest, A. van der, H. van der Kooij, T. van Rossum, W. Ruijssenaars, W. Swüste & T. Westerbeek, *Spellen met spellingzwakke kinderen*. 's-Hertogenbosch: Malmberg, 1980.

Groenendaal, J. H. A., Reactie: taalontwikkelingsstoornissen en leerproblemen. *Tijdschrift voor Orthopedagogiek*, 1987, *26*, 249-253.

Hallahan, D. P. & J. M. Kauffman, *Introduction to learning disabilities*. Englewood Cliffs (N.J.): Prentice-Hall Inc., 1976.

Heesen, H., D. Strelitski & A. van der Wissel, *Schiedamse Rekentest*. Groningen: Tjeenk Willink, 1970.

Johnston, J. & T. K. Shery, The use of grammatical morphemes by children with communication disorders. In: D. Morehead & A, Morehead (Eds.), *Normal and deficient child language*. Baltimore: University Park Press, 1976.

Leij, A. van der, *Ernstige leesproblemen*. Lisse: Swets & Zeitlinger, 1983.

Morais, J., J. Alegria & A. Content, The relationships between segmental analysis and alphabetic literacy: an interactive view. *Cahiers de Psychologie Cognitive*, 1987, *7*, 415-438.

Morehead, D. & D. Ingram, The development of base syntax in normal and linguistically deviant children. *Journal of Speech and Hearing Research*, 1973, *16*, 330-352.

Pauw, L. M. J., Een onderzoek naar taalontwikkelingsstoornissen bij kinderen met leerproblemen. *Tijdschrift voor Orthopedagogiek*, 1987, *26*, 33-48.

Pressman, E., D. Roche, J. Davey & P. Firestone, Patterns of auditory perception skills in children with learning disabilities: a computer-assisted approach. *Journal of Learning Disabilities*, 1986, *19*, 485-488.

Ripich, D. N. & P. L. Griffith, Narrative abilities of children with learning disabilities and nondisabled children: story structure, cohesion, and propositions. *Journal of Learning Disabilities*, 1988, *21*, 165-173.

Santen, A. van, *De morfologie van het Nederlands*. Dordrecht: Foris, 1983.

Schaerlaekens, A. M. & S. Gillis, *De taalverwerving van het kind*. Groningen: Wolters-Noordhoff, 1987.

Spekman, N. J., Dyadic verbal communication abilities of learning disabled children and normally achieving fourth- and fifth-grade boys. *Learning Disability Quarterly*, 1981, *4*, 139-151.

Stanovich, K. E., R. G. Nathan & M. Vala-Rossi, Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. *Reading Research Quarterly*, 1986, *21*, 267283.

Vellutino, F. R., *Dyslexia: theory and research*. Cambridge: MIT Press, 1979.

Vogel, S. A., A qualitative analysis of morphological ability in learning disabled and achieving children. *Journal of Learning Disabilities*, 1983, *16*, 416-420.

Wiig, E. H., E. M. Semel & M. B. Crouse, The use of English morphology by high risk and learning disabled children. *Journal of Learning Disabilities*, 1973, *6*, 457-465.

Williams, F., *Exploration of the linguistic attitudes of teachers*, Rowley (Mass.): Newbury House, 1976.

Williams, F., Psychological correlates of speech characteristics: on sounding disadvantaged. *Journal of Speech and Hearing Research*, 1970, *13*, 472-493.

Wissel, A. van der, *Schooluitval*. Groningen: Academisch Proefschrift, 1984.

Curricula vitae

W. H. J. van Bon (1946), psycholoog, is als universitair hoofddocent verbonden aan de sectie Leerproblemen van de Vakgroep Orthopedagogiek van de KUN. Hij publiceerde onder meer "Taaltests voor Kinderen" (1982) en de bijbehorende validiteitsstudie "Aspecten van Taalvaardigheid" (1984).

M. H. G. Dierx (1963) studeerde Orthopedagogiek aan de KUN met als specialisatie Leerproblemen. Zij werkte mee aan het Invoeringsprogramma Speciaal Onderwijs Nederlandse Taal. Sinds 1987 is zij als orthopedagoge verbonden aan een LOM-school in Nijmegen.

J. J. M. P. Klerkx (1962) studeerde Orthopedagogiek aan de KUN met als specialisatie Leerproblemen. Zij is als medewerkster verbonden aan het onderzoeksproject "De effectiviteit van flashcards" van de KUN-vakgroep Orthopedagogiek.

Adres: Instituut voor Orthopedagogiek, Katholieke Universiteit Nijmegen, Postbus 9103, 6500 HD Nijmegen

Manuscript aanvaard 18-7-'89

Summary

Bon, W. H. J. van, M. H. G. Dierx & J. J. M. P. Klerkx. 'The morphological competence of children with learning disabilities.' *Pedagogische Studiën*, 1990, *67*, 59-69

Seven- and eight-year-old children with learning disabilities appear to be behind in morphological competence as compared with normally achieving children of their age. The retardation is alike for all morphological categories. There is no difference in the kinds of errors which are made.

The retardation does not pertain to the whole domain of cognitive development. The learning disabled have, indeed, lower scores on the Coloured Progressive Matrices, but these cannot account for their low morphological competence.

Other factors than the lower intelligence and the lower morphological competence of the learning disabled are needed for a full explanation of their backwardness in learning.

Kennis en het oplossen van natuurkundige problemen*
Enkele aandachtspunten voor de onderwijspraktijk

M. VANDERLOCHT en J. VAN DAMME
Katholieke Universiteit Leuven
Afdeling Didactiek

Samenvatting

Het uitgangspunt van dit artikel is de overtuiging dat prestaties bij het oplossen van natuurkundige problemen in belangrijke mate bepaald worden door de kwaliteit van het kennisbestand waarover de oplosser beschikt. Er komen vier thema's aan bod die, meer dan in het onderwijs veelal gebruikelijk is, aandacht verdienen als men studenten een adequaat kennisbestand wil bijbrengen. Deze thema's zijn: operationele kennis, selectiekennis, differentiatie tussen verwante concepten en intuïtieve vooropvattingen. Voor elk thema bespreken we eerst de betekenis en de relevantie. Vervolgens komen de resultaten van constaterend en construerend onderzoek aan bod.

1 *Inleiding*

Uit een vorig literatuuroverzicht (Vanderlocht & Van Damme, 1989a) is gebleken dat prestaties bij het oplossen van natuurkundige problemen onder meer bepaald worden door de probleemaanpak die de oplosser volgt. Een oplossingsproces is efficiënter naarmate de oplosser op een systematische wijze te werk gaat. Instructies in probleemaanpak kunnen onder bepaalde voorwaarden leiden tot prestatieverbeteringen. Eén van de voorwaarden is dat de oplosser aan zulke instructies de nodige vak- en probleemspecifieke kenniselementen moet kunnen koppelen. Naast de probleemaanpak zal dus zeker ook de kwaliteit van het kennisbestand waarover de oplosser beschikt, bepalend zijn voor zijn prestaties.

De bedoeling van deze bijdrage is niet een afgerond theoretisch kader te presenteren i.v.m. kennis en het oplossen van natuurkundige problemen. Onze benadering is eerder pragmatisch. We bespreken vier thema's die, meer dan in het onderwijs veelal gebruikelijk is, aandacht verdienen als men studenten een adequaat kennisbestand wil bijbrengen. Deze thema's zijn: operationele kennis, selectiekennis, differentiatie tussen verwante concepten en intuïtieve vooropvattingen. De bespreking van elk thema bevat drie delen. Vooreerst lichten we de betekenis en de relevantie van het thema toe. Vervolgens bespreken we de resultaten van constaterend en construerend onderzoek ter zake. We beogen hierbij geen exhaustief overzicht van de onderzoeksliteratuur. Het is de bedoeling telkens enkele studies aan te halen die het belang van het betreffende thema onderstrepen.

De thema's en de studies zijn geselecteerd in functie van het vak natuurkunde. Nochtans zijn we van mening dat de behandelde thema's in zekere mate ook relevant zijn voor het oplossen van problemen in andere vakgebieden zoals bijvoorbeeld wiskunde, statistiek en scheikunde. We zullen overigens af en toe verwijzen naar onderzoek dat betrekking heeft op één van die domeinen.

2 *Operationele kennis van concepten en principes*

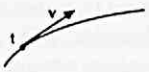
2.1 *Betekenis en relevantie*

Om natuurkundige problemen op te lossen moet een student niet alleen de betekenis kennen van een hele reeks concepten en principes, hij moet deze tevens kunnen toepassen in uiteenlopende situaties. Verschillende auteurs (Anderson, 1982; De Jong & Ferguson-Hessler, 1982) maken in dit verband een onderscheid tussen declaratieve of beschrijven-

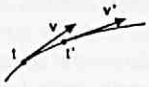
* De literatuurstudie gebeurde in het kader van het ISSAD-project aan het Onderzoekscentrum voor Secundair en Hoger Onderwijs van de Afdeling Didactiek. Het ISSAD-project is een interuniversitair F.K.F.O.-project op initiatief van de Minister van Onderwijs.

Pedagogische Studiën 1990 (67) 70-81

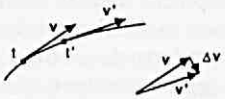
(1) Original velocity (v)
Draw the vector **v** at the time of interest



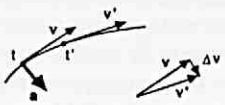
(2) New velocity (v')
Draw the vector **v'** at a slightly later time



(3) Change of velocity (Δv)
Draw a separate vector diagram so that the arrow tails of **v** and **v'** coincide. Construct the vector **Δv** which is the vector drawn from the head of the original velocity **v** to the head of the new velocity **v'**.



(4) Acceleration (a)
Divide the vector **Δv** by Δt to obtain a new vector **Δv**/Δt having the same direction as **Δv** (but different magnitude and units). If the time interval Δt is sufficiently small, this vector is the acceleration **a**.
Draw the vector **a** at the time of interest.



Figuur 1 *Operationele definitie van het concept versnelling (Labudde, Reif & Quinn, 1988, p. 83)*

de kennis en operationele of procedurele kennis. Reif (1985, 1987) maakt een analoog onderscheid tussen een declaratieve en een operationele definitie van een wetenschappelijk concept of principe. In een declaratieve definitie wordt de betekenis van het wetenschappelijk concept of principe verduidelijkt. Een declaratieve definitie van het concept versnelling is **a** = dv/dt. Het toepassen van de grootheid versnelling in concrete situaties impliceert echter een aantal operaties die in zulke declaratieve definitie niet geëxpliciteerd worden. Dit gebeurt wel in een operationele definitie. Een operationele definitie van het concept versnelling volgens Labudde, Reif en Quinn (1988, p. 83) is weergegeven in Figuur 1.

Volgens deze auteurs wordt in het natuurkundeonderwijs nogal eens nagelaten zulke operationele kenniselementen te expliciteren. Ook Eylon en Linn (1988, p. 278) delen deze mening: "Students must have a repertoire of these specific procedures if they are to acquire in-depth understanding of a science topic. Instructors rarely help students develop such a repertoire."

2.2 *Constaterend onderzoek*

Dat studenten moeilijkheden hebben met het toepassen van het concept versnelling blijkt onder meer uit de resultaten op de voormeting in een construerend onderzoek van Labudde et al. (1988). Aan het onderzoek werd deelgenomen door zes eerstejaars universiteitsstudenten die een inleidende cursus natuurkunde volgden. In de voormeting werden vijf bewegingssituaties voorgelegd (o.a. een rechtlijnige beweging met ogenblikkelijke snelheid nul en een kromlijnige beweging met constante snelheid). De studenten dienden telkens aan te geven of de versnelling al dan niet nul was en moesten tevens de richting van de versnelling met een pijl aanduiden. Tweeëntwintig van de 30 antwoorden over de grootte en 12 van de 30 antwoorden over de richting waren correct. Slechts 40% van de antwoorden waren correct zowel wat grootte als richting betreft. Typische fouten zijn de opvattingen dat de versnelling nul is bij een cirkelvormige beweging met constante snelheid en bij een trillende veer op het laagste punt. Op basis van de uitspraken van de studenten tijdens het oplossingsproces werd een inventaris gemaakt van de inductieregels die zij vermeldden of toepasten. In totaal werd 21 keer een correcte inductieregel vastgesteld. Een correcte inductieregel is bijvoorbeeld: als de snelheid afneemt of toeneemt is de versnelling verschillend van nul. Drieëndertig keer werd een foutieve inductieregel vastgesteld. Een voorbeeld van een fou-

tieve inductieregel is: als de grootte van de snelheid constant is, is de versnelling nul.

Ferguson-Hessler en De Jong (1987) lieten een groep eerstejaarsstudenten van een Technische Universiteit een tekst over een natuurkundig onderwerp hardop denkend bestuderen. Ze stelden vast dat 'zwakke studenten' zich bij de studie vooral concentreerden op de declaratieve kennis terwijl 'goede studenten' ook aandacht hadden voor de operationele kennis die in de tekst aan bod kwam.

Een onderzoek van Linn, Sloane en Clancy (1987) heeft betrekking op het leren programmeren in Pascal. Zij onderzochten op een steekproef van 14 klassen het verband tussen enkele kenmerken van het gegeven onderwijs en de bereikte programmeervaardigheid van de studenten. Er werd o.m. vastgesteld dat de bereikte programmeervaardigheid beter was in klassen waar bepaalde procedures voor het ontwerpen, implementeren, testen en 'debuggen' van programma's expliciet onderwezen werden.

2.3 *Construerend onderzoek*

In deze paragraaf gaan we verder in op het construerend onderzoek van Labudde et al. (1988). De resultaten op de voormeting in dit onderzoek werden hiervoor besproken. Een aansluitende experimentele sessie had tot doelstelling de studenten een adequate en coherente operationele kennis van de grootheid versnelling bij te brengen. Hierbij werd gebruik gemaakt van de operationele definitie die in Figuur 1 weergegeven is. In een eerste fase werd deze operationele definitie stap voor stap toegelicht, geïllustreerd en ingeoefend op enkele voorbeelden. In een tweede fase werd de operationele definitie gebruikt bij het evalueren van eigen en andermans antwoorden. Elke student werd geconfronteerd met zijn eigen antwoorden op de voormeting en gevraagd deze te evalueren door toepassing van de operationele definitie. Indien het antwoord op een vraag uit de voormeting foutief bleek te zijn, diende de student tevens een verklaring te zoeken voor zijn fout. Vervolgens werden twee probleemsituaties en de antwoorden erop van een 'hypothetische oplosser' voorgelegd. Aan de studenten werd gevraagd deze antwoorden te evalueren, er een mogelijke verklaring voor te zoeken en ten slotte zelf het goede antwoord te geven. De studenten slaagden er vlot in om de operationele definitie toe te passen in deze leertaken. Vijftien van de 18 eigen fouten en 34 van de 36 aangeboden hypothetische antwoorden werden adequaat geëvalueerd en verklaard. Tachtig procent van de foute inductieregels die in de voormeting werden vastgesteld, werden bij het oplossen van deze taken expliciet gecorrigeerd door de studenten.

In de nameting werden dezelfde vijf bewegingsvormen als in de voormeting gebruikt; ze werden echter wel ingekaderd in andere probleemsituaties. Een eerste vaststelling is dat de studenten tijdens de nameting vrij frequent 'spontaan' de operationele definitie gebruiken. In 18 van de 30 oplossingen werd de operationele definitie direct toegepast en in vier van de 30 oplossingen werd de operationele definitie gebruikt om een eerder gegeven antwoord te controleren. De studenten bleken in de nameting bovendien beduidend beter te presteren dan in de voormeting. Het aantal vragen dat zowel wat grootte als wat richting betreft correct werd opgelost, steeg van 40% naar 95%. Tijdens het oplossen van de opgaven werden bijna geen foute inductieregels vermeld.

3 *Selectiekennis*

3.1 *Betekenis en relevantie*

Wanneer aan een student een natuurkundig probleem voorgelegd wordt, dan zal hij op basis van kenmerken van de probleemsituatie dienen uit te maken met behulp van welke wet of formule hij het probleem gaat oplossen. Dit selectieproces zal efficiënter verlopen als de student voor elk van de beschikbare wetten en formules een beeld heeft van de kenmerken die in de probleemsituatie aanwezig moeten zijn opdat de betreffende wet of formule bruikbaar zou zijn. Hij dient m.a.w. een inzicht te hebben in de toepassingsmogelijkheden van verschillende wetten en formules. De Jong en Ferguson-Hessler (1982) geven hiervan een mooi voorbeeld uit het vakgebied 'Electriciteit & Magnetisme'. In dit vakgebied moet vaak met 'de wet van Gauss' gewerkt worden. Deze wet heeft twee vormen, één voor homogene en één voor niet-homogene velden. Het kenmerk homogeen of niet-homogeen zal dus bepalend zijn voor de te selecteren vorm van de wet.

Studenten dienen niet alleen inzicht te heb-

ben in de toepassingsmogelijkheden van wetten en formules; zij dienen tevens de toepassingsbeperkingen ervan te kennen. Vele natuurkundige wetten en formules kunnen slechts onder welbepaalde voorwaarden afgeleid of vastgesteld worden. Ze mogen bijgevolg ook enkel toegepast worden in situaties die voldoen aan deze voorwaarden. Zo kan bij een projectielbeweging de reikwijdte, d.w.z. de horizontale afstand tussen het vertrekpunt en de plaats waar het projectiel op de grond terecht komt, berekend worden met de formule $R = v(0)^2 . \sin 2\Theta(0)/g$. Deze formule is echter enkel geldig wanneer de lanceerhoogte gelijk is aan de eindhoogte. Ze geldt bijvoorbeeld niet wanneer een steen vanaf 1 m hoogte schuin wordt weggeworpen en een eind verder op de grond terecht komt.

Omdat kennis van toepassingsmogelijkheden en -beperkingen een belangrijke rol speelt bij de selectie van een wet of formule ter oplossing van een bepaald probleem, kunnen we in navolging van De Jong en Ferguson-Hessler (1982) spreken van selectiekennis.

3.2 *Constaterend onderzoek*

Het toepassen van wetten of formules in situaties die niet voldoen aan de toepassingsvoorwaarden is een belangrijke bron van fouten bij het oplossen van natuurkundige problemen. We hebben dit zelf vastgesteld bij een analyse van 152 oplossingen van problemen over de projectielbeweging (Vanderlocht & Van Damme, 1989b). Bij de behandeling van dit leerstofonderdeel in de cursus van de proefgroep werden naast de algemene formules die het bewegingsverloop beschrijven een drietal supplementaire formules afgeleid om op een meer directe manier de reikwijdte (R), de vluchttijd (T) en de maximale hoogte (H) te berekenen. Zoals reeds vermeld, mogen deze formules enkel toegepast worden als de lanceerhoogte gelijk is aan de eindhoogte. Slechts één van de zeven opgaven die in het onderzoek werden opgenomen, voldeed aan deze voorwaarde. We stelden vast dat 39 keer een RTH-formule ongeoorloofd werd toegepast. Twintig studenten (13,2%) pasten één of meerdere RTH-formules ongeoorloofd toe.

Het is mogelijk dat een aantal van deze studenten de genoemde toepassingsbeperking van de RTH-formules wel kende maar ze ten gevolge van een impulsieve werkwijze uit het oog verloor. We hadden echter de indruk dat een groot deel van de studenten de toepassingsbeperking gewoon niet kende. In een vervolgonderzoek hebben we dit getoetst (Vanderlocht, 1989). Aan 27 studenten werd een probleem over de projectielbeweging voorgelegd waarin de lanceerhoogte verschilt van de eindhoogte. Aan de studenten werd gevraagd of de formule voor de reikwijdte in deze opgave kon gebruikt worden; ze moesten hun antwoord ook motiveren. Van de 27 studenten gaven er 23 een nee-antwoord en 4 een ja-antwoord. Vijftien studenten (55,5%) gaven in hun verantwoording op één of andere manier aan dat de formule voor de reikwijdte niet mag gebruikt worden omdat de lanceerhoogte verschilt van de eindhoogte. Twaalf studenten (45,5%) slaagden er niet in om de relevante toepassingsvoorwaarde van de formule voor de reikwijdte te verwoorden.

Een onderzoek van Allwood en Montgommery (1981) heeft betrekking op het oplossen van statistische problemen. Twintig studenten dienden elk twee statistische problemen op te lossen. In totaal werden 82 fouten geïdentificeerd in de oplossingen. Om na te gaan in welke mate kennistekorten aan de basis liggen van deze fouten dienden de studenten in eerste instantie hun oplossingen uit te leggen en te verantwoorden. In tweede instantie stelden de onderzoekers zelf gerichte kennisvragen in aansluiting bij de geïdentificeerde fouten. Bij 63 van de 82 fouten (77%) bleken de antwoorden van de studenten aanwijsbare kennistekorten te bevatten. Negentien keer werd een tekort in kennis van toepassingsvoorwaarden van statistische formules vastgesteld. Elf keer bleek zulke kennis van toepassingsvoorwaarden te ontbreken en acht keer werden onjuiste toepassingsvoorwaarden vermeld.

Chi, Feltovich en Glaser (1981) legden aan beginners en experts in het oplossen van natuurkundige problemen een reeks opgaven voor met de opdracht deze te sorteren in groepen van gelijksoortige opgaven. De beginnende oplossers bleken zich bij het sorteren te laten leiden door 'oppervlaktekenmerken'. Ze groepeerden opgaven waarin dezelfde objecten, situatiekenmerken of letterlijke termen voorkomen. De experts daarentegen sorteerden de opgaven spontaan volgens het natuurkundige principe waarmee het probleem kan opgelost worden. Gelijkaardige resultaten werden vastgesteld bij het sorteren van wiskundige (Schoenfeld & Herrmann, 1982) en

statistische problemen (Been & Brokken, 1986).

Aan de reeds genoemde proefpersonen legden Chi et al. (1981) tevens enkele natuurkundige problemen voor met de vraag welke 'basic approach' ze zouden volgen bij de oplossing ervan. Uit de hardop denken protocollen bleek onder meer dat de experts in dit selectieproces beroep doen op bepaalde selectiecriteria om uit te maken of een bepaald principe bruikbaar is. De auteurs formuleren het als volgt (Chi et al., 1981, p. 147): "The analysis of features suggests that experts perceive more in a problem statement than do novices. They have a great deal of tacit knowledge that can be used to make derivations from the situation described by the problem statement. Their selection of the 'principle' to apply seems to be guided by this second order, derived knowledge."

Niet alleen experts en beginners, ook goede en zwakke beginnende oplossers blijken verschillend te presteren bij het selecteren van een wet of formule ter oplossing van voorgelegde problemen. Uit onderzoek van De Jong (1986, pp. 131-165) blijkt dat het kunnen selecteren van relevante kenniselementen bij voorgelegde problemen in sterke mate verband houdt met prestaties bij het oplossen van problemen. Eerstejaarsstudenten van een Technische Universiteit kregen een reeks opgaven over Electriciteit en Magnetisme aangeboden en zij dienden aan te geven welke kenniselementen volgens hen bruikbaar zijn voor de oplossing van elke opgave. Goede oplossers bleken duidelijk beter dan zwakke oplossers in staat om aan een probleemsituatie de vereiste kenniselementen te koppelen. De correlatie tussen het tentamencijfer en de hoeveelheid gegeven van toepassing zijnde kenniselementen was 0,77.

3.3 *Construerend onderzoek*

We hebben in de literatuur geen construerend onderzoek gevonden dat specifiek gericht is op het expliciteren van toepassingsbeperkingen. We kunnen wel verwijzen naar een onderzoek van Reif, Larkin en Brackett (1976). Zij onderwezen studenten een algemene strategie om natuurkundige formules te verwerken. De studenten werden gestimuleerd om bij de confrontatie met een nieuwe formule een reeks van inzichtbevorderende vragen te doorlopen. Ze dienden onder meer telkens de toepassingsbeperkingen van de formule te expliciteren en deze te illustreren met concrete probleemsituaties. Deze algemene leerstofverwerkingsstrategie bleek goede resultaten op te leveren. Specifieke resultaten over de selectiekennis en het selectieproces worden echter niet vermeld.

Alexander en Judy (1988) verwijzen naar een onderzoek van Dellarosa (1984) waaruit blijkt dat de selectiekennis van leerlingen kan bevorderd worden door hen gerichte vergelijkingen te doen maken tussen verschillende opgaven. In een eerste experiment dienden verschillende groepen leerlingen een reeks algebraïsche woordproblemen te lezen en ze kregen hierbij de opdracht te antwoorden op voorgelegde vragen. De aard van de vragen was verschillend per groep. Sommige vragen richtten de aandacht vooral op kenmerken van de afzonderlijke opgaven terwijl andere vragen de leerlingen stimuleerden gerichte vergelijkingen te maken tussen de verschillende opgaven. Na het lezen van de opgaven werd aan de leerlingen gevraagd de opgaven te sorteren in groepjes. Men stelde vast dat de leerlingen die door de vragen gestimuleerd werden gerichte vergelijkingen te maken tussen de verschillende opgaven, deze sorteerden op gelijkaardige wijze als experts. In een vervolgexperiment bleek dat leerlingen na zulke interventie er beter in slaagden om aan een reeks opgaven de gepaste algebraïsche representatie te koppelen.

4 *Differentiatie tussen verwante concepten*

4.1 *Betekenis en relevantie*

Bouwstenen van de natuurkunde zijn concepten van een bijzondere aard, namelijk grootheden. Een grootheid is een variabele waarmee men op kwantificeerbare wijze een kenmerk van een object of een situatie kan beschrijven. Naast een beperkt aantal basisgrootheden die op onafhankelijke wijze gedefinieerd zijn, worden de meeste grootheden gedefinieerd via een rekenkundige relatie met andere grootheden. Grootheden zoals bijvoorbeeld positie, snelheid en versnelling zijn duidelijk verwant met elkaar maar hebben toch een eigen specifieke betekenis. Een correct begrip en gebruik veronderstellen dan ook dat men zulke verwante grootheden uit el-

kaar kan houden. Onderzoek toont echter aan dat studenten verwante grootheden meermaals met elkaar verwarren.

4.2 *Constaterend onderzoek*

In een onderzoek van Trowbridge en McDermott (1980) dienden studenten het concept 'snelheid' toe te passen bij de beoordeling van reële bewegingssituaties. Hierbij werd gebruik gemaakt van 'snelheidvergelijkingstaken'. In zulke taak ziet de student twee knikkers gelijktijdig rollen over parallel opgestelde U-vormige buizen. Aan de studenten wordt gevraagd of de twee knikkers ergens dezelfde snelheid hebben. Een behoorlijk aantal studenten antwoordt bij deze taak dat de twee knikkers dezelfde snelheid hebben op de momenten dat ze elkaar voorbijsteken. Op die momenten hebben de knikkers dezelfde positie, echter niet dezelfde snelheid. De auteurs spreken dan ook van een positie-snelheid verwarring. Om te controleren of deze positie-snelheid verwarring niet grotendeels te wijten was aan de perceptuele dominantie van het voorbijsteken, werd een tweede snelheidvergelijkingstaak afgenomen. In deze taak hebben de knikkers op één bepaald moment dezelfde snelheid maar ze steken elkaar niet voorbij. Ook bij deze taak wijzen de antwoorden van verscheidene studenten op een positie-snelheid verwarring. Deze studenten argumenteren dat de knikkers nergens dezelfde snelheid hebben omdat ze nergens naast elkaar komen. Bij de aanvang van het vak fysica lag het percentage studenten dat deze taken correct oplost, afhankelijk van de groep, tussen 40 en 70. Na het volgen van het vak natuurkunde lagen de percentages tussen 70 en 90. Bijna alle fouten waren positie-snelheid verwarringen.

In de lijn van dit onderzoek heeft Hewson (1985) een computergestuurde procedure ontwikkeld waarbij studenten dienen te beoordelen wanneer twee op een scherm voortbewegende auto's dezelfde snelheid hebben. Ook met deze procedure stelde Hewson bij verschillende groepen studenten vast dat de positie-snelheid verwarring vrij frequent voorkomt.

Een tweede onderzoek van Trowbridge en McDermott (1981) heeft betrekking op het concept versnelling. Hierbij werd gebruik gemaakt van een analoge proefopstelling als die in het onderzoek van 1980. Deze keer werd aan de studenten gevraagd of de knikkers dezelfde of een verschillende versnelling hebben. Analoog aan de juist vermelde positie-snelheid verwarring kwam bij deze taak een positie-versnelling verwarring te voorschijn. Meerdere studenten antwoorden dat de knikkers dezelfde versnelling hebben op de momenten dat ze elkaar voorbijsteken. Andere studenten verwarren dan weer de concepten snelheid en versnelling. Zij stellen dat één van de knikkers een grotere versnelling heeft omdat hij de andere knikker nadert en dus sneller rolt.

Zulke gebrekkige conceptdifferentiatie speelt studenten ook parten bij het oplossen van kwantitatieve vraagstukken. Trowbridge en McDermott (1981) legden studenten een vraagstuk voor waarin de versnelling van een knikker die van een helling rolt, moet berekend worden. Na het volgen van het vak natuurkunde losten, afhankelijk van de groep, tussen 18% en 28% van de studenten dit vraagstuk correct op. Het grootste deel van de fouten bestond erin dat men bij het berekenen van de versnelling geen onderscheid maakt tussen enerzijds de gemiddelde snelheid tijdens de afdaling en anderzijds het snelheidsverschil tussen het einde en het begin van de afdaling.

4.3 *Construerend onderzoek*

Reif (1985) pleit ervoor om bij het aanbrengen van een natuurkundige grootheid expliciet aandacht te schenken aan de relatie met en het onderscheid tussen verwante grootheden, zowel op het niveau van de definiëring als met behulp van goed gekozen voorbeeldsituaties.

Deze didactische wenk zien we gerealiseerd bij Rosenquist en McDermott (1987) die een speciale onderwijsmodule voor het leerstofonderdeel kinematica ontwikkelden. Volgens deze auteurs wordt in de meeste natuurkunde-lessen sterk de nadruk gelegd op het algebraïsche formalisme waarmee natuurkundige grootheden en wetten kunnen uitgedrukt worden en wordt er te weinig aandacht besteed aan het verduidelijken van de concepten die aan zulke formalismen ten grondslag liggen. De onderwijsmodule 'kinematica' legt sterk de nadruk op conceptvorming. Een belangrijk onderdeel van de module heeft als doel de eigenheid van en het verschil tussen de verwante kinematische grootheden positie, snelheid en versnelling bij te brengen.

Om het onderscheid tussen de concepten positie en snelheid te verduidelijken worden de snelheidvergelijkingstaken uit het onderzoek van Trowbridge en McDermott (1980) voorgelegd. De bedoeling van een eerste taak, waarbij de twee knikkers elkaar voorbijsteken, is bij studenten die de concepten positie en snelheid slechts vaag onderscheiden een typische positie-snelheid beoordelingsfout uit te lokken. Zulke fout vormt het uitgangspunt om studenten uit te leggen en te demonstreren dat objecten dezelfde snelheid kunnen hebben zonder zich zij aan zij te bevinden en dat een object een ander object maar kan voorbijsteken als het een grotere snelheid heeft.

Het onderscheid snelheid-versnelling wordt in twee stappen aangebracht. In een eerste experiment laat men herhaaldelijk een knikker van een helling rollen. Boven de helling is een gemarkeerde band opgesteld waarvan de snelheid instelbaar is. De snelheid van de band wordt gebruikt als vergelijkingsstandaard. De studenten dienen te meten hoelang het duurt alvorens de knikker dezelfde snelheid heeft als de band. De bedoeling van dit experiment is ten eerste de studenten te doen inzien dat snelheid een continu veranderende grootheid is met een concreet maatgetal op elk tijdstip. Zo wordt tevens het onderscheid tussen snelheid en snelheidsverschil manifest. Het meten van de tijdspanne die de knikker nodig heeft om de snelheid van de band te bereiken, maakt de studenten bewust van het tijdsinterval waarin een snelheidsverandering plaatsvindt en zo zijn de voorwaarden gecreëerd om het concept versnelling te begrijpen als $\Delta v/\Delta t$. Na dit voorbereidend experiment wordt de versnellingvergelijkingstaak voorgelegd. Bij vage conceptdifferentiatie zal deze taak typische beoordelingsfouten uitlokken. In aansluiting bij zulke fouten start de leerkracht een onderwijsleergesprek waarin de noties ogenblikkelijke snelheid, snelheidsverschil, corresponderend tijdsinterval en uiteindelijk versnelling duidelijk worden onderscheiden en geïllustreerd aan de hand van de beoordelingstaak.

De experimentele module bleek duidelijk resultaten op te leveren. Op de snelheidvergelijkingstaken steeg het percentage succesvolle studenten van 60 naar 90 en op de versnellingvergelijkingstaak werd een stijging vastgesteld van 0 naar 40. Bij een andere groep studenten die een meer wiskundig onderbouwde cursus natuurkunde volgden, was de stijging veel geringer. Het percentage succesvolle snelheidvergelijkers bleef constant op 60 en het percentage succesvolle versnellingvergelijkers steeg van 20 naar 40.

5 *Intuïtieve vooropvattingen*

5.1 *Betekenis en relevantie*

Studenten verschijnen niet als een 'tabula rasa' in de natuurkundelessen. Verschillende concepten die in de natuurkunde strikt gedefinieerd zijn, worden ook in de dagelijkse taal gebruikt, zij het met een minder afgelijnde en soms afwijkende betekenis. Neem bijvoorbeeld het concept 'arbeid'. Iedereen weet dat het heel wat moeite kost om een zwaar voorwerp op een zekere hoogte vast te houden. Natuurkundigen beweren nochtans dat de kracht die deze persoon uitoefent op het voorwerp geen 'arbeid' levert. Verder blijken studenten reeds voor het volgen van onderwijs in de natuurkunde bepaalde intuïtieve opvattingen te hanteren om natuurkundige verschijnselen te verklaren. Lijnse (1982) maakt in dit verband een onderscheid tussen een 'straatbeeld' en een 'schoolbeeld' van natuurkundige fenomenen. Met de term straatbeeld verwijst Lijnse naar zulke intuïtieve vooropvattingen. Een straatbeeld is volgens Lijnse sterk geworteld in directe ervaringen, bestaat uit vage begrippen en onduidelijke relaties, bevat soms logische inconsistenties, is het resultaat van een individuele momentane zingeving en heeft een intuïtief karakter. Met de term schoolbeeld verwijst Lijnse naar de wetenschappelijke beschrijving en verklaring van natuurkundige fenomenen zoals die voorkomen in handboeken en op school onderwezen worden. Dit schoolbeeld is veel minder geworteld in directe ervaringen, bestaat uit scherp afgelijnde begrippen en precieze relaties en is logisch consistent. Verder gaat het om publiek gegeneraliseerde en blijvende kennis die het resultaat is van een expliciete reflectie.

Niettegenstaande intuïtieve vooropvattingen in vele gevallen afwijken van de wetenschappelijke verklaringen, blijken ze ook na het volgen van natuurkunde-onderwijs in sterke mate te persisteren en een diepgaand begrip van de wetenschappelijke verklaringen te bemoeilijken. Deze vaststelling stimuleerde meerdere onderzoekers om onderwijsstrate-

gieën te ontwikkelen waarin ten volle rekening gehouden wordt met de aanwezigheid van intuïtieve vooropvattingen.

5.2 *Constaterend onderzoek*

De laatste jaren is reeds heel wat onderzoek verricht naar de vooropvattingen van studenten over de meest uiteenlopende fenomenen die in de verschillende natuurwetenschappen aan bod komen. Pfundt en Duit (1988) geven een bibliografisch overzicht dat liefst 188 bladzijden beslaat. Het is dan ook niet onze bedoeling hier een overzicht te schetsen van dit uitgebreide onderzoeksterrein. We zullen ons beperken tot het beschrijven van enkele vooropvattingen over concepten en fenomenen uit het domein van de mechanica.

De grootheid kracht wordt in de dynamica gebruikt om interacties tussen objecten te beschrijven. In het dagelijks taalgebruik heeft deze term een aantal connotaties die afwijken van de natuurkundige betekenis. Zo doet 'kracht' vooral denken aan een actieve gebeurtenis zoals bijvoorbeeld het opheffen van een voorwerp en wordt het in veel mindere mate geassocieerd met interacties waarbij een beweging wordt tegengewerkt. Minstrell (1982) vroeg aan een groep leerlingen welke krachten inwerken op een boek dat stilligt op tafel. Veertien op 27 leerlingen vermelden enkel een neerwaartse kracht. De kracht van de tafel op het boek, die ervoor zorgt dat het boek niet naar beneden valt, wordt blijkbaar spontaan niet als dusdanig geconcipieerd. Waar kracht als natuurkundig begrip verwijst naar een interactie tussen objecten, wordt het in de dagelijkse taal nogal eens beschouwd als een kenmerk van en in een object. Wanneer bijvoorbeeld een bodybuilder gewichten heft, spreken leerlingen spontaan over de kracht van de bodybuilder en niet van de kracht die de bodybuilder uitoefent op de gewichten (Vegting, 1985).

McCloskey en zijn medewerkers (McCloskey, 1983; Green, McCloskey & Caramazza, 1985) beschrijven een reeks interessante studies over het voorspellen en verklaren van bewegingen. Zo legden zij aan verschillende groepen studenten een reeks opdrachten voor waarbij een voorwerp doorheen een horizontaal gelegen, curvevormige buis voortgestuwd wordt. De vraag luidt telkens welke baan het voorwerp zal volgen als het de buis verlaat. Verscheidene studenten beweren dat het voorwerp een curvevormig traject zal volgen. Bij studenten die geen cursus natuurkunde volgden, gaf 50% zulk antwoord. Bij studenten die een cursus natuurkunde op high school of college niveau volgden, gaven respectievelijk 33% en 12% van de studenten zulk antwoord. Meerdere studenten gaven als verklaring dat het voorwerp bij het voortstuwen een soort inwendige kracht meekrijgt. Die inwendige kracht zorgt ervoor dat het voorwerp na het verlaten van de buis zijn curvevormige baan verder zet. Volgens de studenten zal deze inwendige kracht geleidelijk afnemen en hierdoor zal ook de curvevorm van de baan geleidelijk verminderen.

Aan een groep van 135 eerstejaars psychologiestudenten vroeg McCloskey (1983) de baan te tekenen van een metalen bol die vanop een rots horizontaal wordt weggeslagen. Vierenzeventig procent van de studenten tekende een min of meer parabolische baan. Tweeëntwintig procent van de studenten tekende een baan waarbij de bol na zekere tijd recht naar beneden valt. Ook hier poneren meerdere studenten dat de bol bij het wegslaan van de rots een soort inwendige kracht meekrijgt die ervoor zorgt dat de bol na het verlaten van de rots nog een tijd horizontaal verder beweegt. Geleidelijk zal deze inwendige kracht 'opgebruikt' geraken, de zwaartekracht krijgt de bovenhand en de bol valt, volgens sommige studenten plots en volgens andere geleidelijk, naar beneden.

Deze typische antwoorden zijn volgens McCloskey uitingen van een achterliggend intuïtief verklaringsschema, door hem impetusopvatting genoemd. Deze impetusopvatting, waarvan meerdere varianten voorkomen, stelt in essentie dat een voorwerp bij het in beweging brengen een zekere inwendige stuwkracht of 'impetus' meekrijgt die tijdens de beweging wordt opgebruikt zodat het voorwerp geleidelijk tot stilstand komt. Deze opvatting wijkt op een belangrijk punt af van Newtons bewegingswetten. In de impetusopvatting worden beweging en kracht geassocieerd terwijl de tweede wet van Newton de grootheden kracht en versnelling associeert.

In dit opzicht vertoont de impetusopvatting duidelijke overeenkomsten met de vaststelling van Clement (1983) dat wij spontaan veronderstellen dat elke beweging de werking van een kracht impliceert. In de bewegingsfenomenen waarmee wij dagelijks ge-

confronteerd worden, spelen wrijvingskrachten een belangrijke rol en in zulke situaties is inderdaad een kracht nodig om een voorwerp in beweging te houden. De opvatting dat elke beweging de werking van een kracht veronderstelt, is dus niet zo verwonderlijk, maar speelt wel een bemoeilijkende rol bij het begrijpen van natuurkundige verklaringen van bewegingssituaties. Clement (1983) vroeg aan een groep ingenieurstudenten welke krachten tijdens de vlucht inwerken op een muntstuk dat verticaal omhoog geworpen wordt als de luchtweerstand buiten beschouwing blijft. Bij de aanvang van het vak natuurkunde gaven 88% van de studenten een fout antwoord. In bijna alle foute antwoorden was naast een neerwaartse pijl ter verwijzing naar de zwaartekracht een opwaartse pijl te vinden. Na het volgen van het vak natuurkunde daalde het percentage foute antwoorden enigszins maar bleef toch verbazend hoog, namelijk 72.

5.3 *Construerend onderzoek*

De vaststelling dat intuïtieve vooropvattingen een bemoeilijkende rol spelen bij het begrijpen van concepten en principes, vormde een nieuwe stimulans en invalshoek voor het ontwikkelen van aangepaste onderwijsmethoden voor de natuurkunde en de natuurwetenschappen in het algemeen (Posner, Strike, Hewson & Gertzog, 1982; Nussbaum & Novick, 1982; Hewson & Hewson, 1983; Joshua & Dupin, 1987).

We vatten enkele aanbevelingen samen die vanuit deze hoek naar voren worden geschoven. Een eerste belangrijke voorwaarde om vat te krijgen op intuïtieve vooropvattingen is deze te expliciteren, m.a.w. studenten bewust maken van hun eigen vooropvattingen. Vervolgens kan men de studenten confronteren met empirische gegevens die niet kloppen met hun vooropvattingen en tevens hun eigen opvattingen vergelijken met de wetenschappelijke verklaringen. Op deze wijze kan een 'conceptueel conflict' gecreëerd worden dat een gunstige voedingsbodem is voor een aanpassing van de eigen opvattingen.

Ervaringen van docenten en onderzoeksresultaten bevestigen de waarde van onderwijs waarin wordt uitgegaan van de vooropvattingen die bij studenten aanwezig zijn.

Hewson en Hewson (1983) beschrijven een onderzoek over het onderwijzen van de concepten massa, volume en densiteit. Zij gebruikten een toets waarin 30 uitspraken over de verschijnselen 'drijven en zinken' moeten beoordeeld worden als voor- en nameting in een construerend onderzoek. Een experimentele groep kreeg een reeks lessen waarbij werd uitgegaan van de vastgestelde vooropvattingen om de juiste wetenschappelijke betekenis van de genoemde concepten te verduidelijken. Een controlegroep kreeg de leerstof aangeboden op de 'traditionele' manier. De experimentele groep presteerde beduidend beter dan de controlegroep op de nameting.

We hebben eerder reeds gewezen op de vaststelling van Minstrell (1982) dat kracht spontaan geassocieerd wordt met een actieve gebeurtenis en in veel mindere mate met het ondersteunen van een object of het tegenwerken van een beweging. Bij de aanvang van een les natuurkunde stelden 13 op 27 leerlingen dat een tafel geen kracht uitoefent op een boek dat stilligt op die tafel. In het vervolg van de les presenteerde Minstrell een reeks andere situaties waarbij het boek in rusttoestand is. Zo werd het boek onder meer op de hand van een leerling gelegd en aan een veer gehangen. De leerlingen dienden telkens aan te geven welke krachten inwerken op het boek en werden gestimuleerd om hun antwoord voor de groep te motiveren. In deze situaties zagen de meeste leerlingen wel in dat er een opwaartse kracht werkzaam is. Aansluitend werd uitvoerig uitgelegd dat het natuurkundige begrip kracht niet enkel slaat op een actieve gebeurtenis zoals stoten of trekken maar ook op situaties zoals het ondersteunen van een object of het tegenwerken van een beweging. Op het einde van de les bleken 25 van de 27 leerlingen ook een opwaartse kracht te vermelden in de situatie waar het boek op een tafel ligt.

Verschillende auteurs (Di Sessa, 1982; White, 1984; Brna, 1987; Van 't Hul, Kulik, Lijnse & Snijders Blok, 1987) ontwikkelden softwarepakketten waarin Newtons bewegingswetten gesimuleerd worden. Studenten kunnen door het indrukken van een knop als het ware een kracht uitoefenen op een voorwerp en zo de bewegingstoestand ervan beïnvloeden. De studenten krijgen de opdracht het voorwerp een aantal opgegeven banen te doen volgen. Op deze speelse wijze worden studenten uitdrukkelijk geconfronteerd met hun eigen vooropvattingen en ontdekken ze de juiste betekenis van Newtons bewegingswetten. White (1984) stelde vast dat studenten na

het doorlopen van zulke opgaven beter scoorden op een toets met toepassingen op de wetten van Newton.

Joshua en Dupin (1987) merken op dat het confronteren van studenten met bepaalde feiten die hun vooropvattingen tegenspreken niet altijd volstaat om hun opvattingen te wijzigen. Studenten zijn zeer creatief in het inpassen van bepaalde gegevens in hun eigen opvattingen. De auteurs wijzen op de mogelijkheden van een goed gekozen analogie om een begripsverandering te ondersteunen. In hun onderzoek gebruikten zij met succes een bepaalde 'treinanalogie' om een correct begrip van bepaalde concepten en principes over electrische schakelingen bij te brengen.

Nussbaum en Novick (1982) wijzen erop dat onderwijs waarbij wordt ingegaan op de vooropvattingen van studenten niet enkel bevorderlijk is voor de begripsvorming maar tevens hun interesse en motivatie voor de lessen sterk stimuleert.

6 *Conclusies*

Tot slot vatten we de geformuleerde aanbevelingen voor het bijbrengen van een adequaat kennisbestand nog even samen.

Ten eerste werd erop gewezen dat voor het oplossen van natuurkundige problemen het niet volstaat enkel de betekenis van concepten en principes te kennen. Het kunnen toepassen van deze concepten en principes in verschillende situaties veronderstelt de kennis van een reeks handelingen die in een declaratieve definitie niet aangegeven worden. Er werd dan ook gepleit voor het expliciteren van operationele kenniselementen.

Om de efficiëntie van het selectieproces bij het oplossen van natuurkundige problemen te bevorderen, is het aangewezen studenten te wijzen op de toepassingsmogelijkheden en -beperkingen van wetten en formules.

Ten derde werd ervoor gepleit om bij het onderwijzen van natuurkundige concepten uitdrukkelijk aandacht te schenken aan de relatie en het onderscheid tussen verwante concepten, zowel op het niveau van de definiëring als met behulp van goed gekozen voorbeeldsituaties.

Ten slotte stelden we vast dat intuïtieve vooropvattingen een bemoeilijkende rol kunnen spelen bij het begrijpen van natuurkundige concepten en principes. Onderzoeksresultaten bevestigen de waarde van onderwijs waarin wordt uitgegaan van de vooropvattingen die bij studenten aanwezig zijn.

Literatuur

Alexander, P.A. & J.E. Judy, The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, 1988, *58*, 375-404.

Allwood, C.M. & H. Montgommery, Knowledge and technique in statistical problem solving. *European Journal of Science Education*, 1981, *4*, 431-450.

Anderson, J.R., Acquisition of cognitive skill. *Psychological Review*, 1982, *89*, 369-406.

Been, P.H. & F.B. Brokken, Kategorisering van statistiekproblemen door beginners en experts. *Tijdschrift voor Onderwijsresearch*, 1986, *11*, 23-34.

Brna, P., Confronting dynamics misconceptions, *Instructional Science*, 1987, *16*, 351-379.

Chi, M.T., H. Feltovich & R. Glaser, Categorisation and representation of physics problems by experts and novices. *Cognitive Science*, 1981, *5*, 121-152.

Clement, J., A conceptual model discussed by Galileo and used intuïtively by physics students. In: D. Getner & L.A. Stevens (Eds.), *Mental models*. Hillsdale: Lawrence Erlbaum, 1983, 325-340.

Di Sessa, A., Unlearning Aristotelian Physics: A study of knowledge based learning, *Cognitive Science*, 1982, *6*, 37-75.

Dellarosa, D., *The role of text comprehension processes and analogical reasoning in the development of problem-solving expertise*. Niet gepubliceerd proefschrift. Boulder, University of Colorado, 1984.

Eylon, B. & M.C. Linn, Learning and instruction: An examination of four research perspectives in science education, *Review of Educational Research*, 1988, *58*, 251-301.

Ferguson-Hessler, M.G.M. & T. de Jong, *Kennisverwerving uit natuurkundige teksten. Een onderzoek naar bestuderingsprocessen van beginnende natuurkundestudenten*. Onderzoekscentrum voor Communicatie en Technische Kennis Overdracht, T.U. Eindhoven, 1987.

Green, B.F., M. McCloskey & A. Caramazza, The relationship of knowledge to problem solving with examples from kinematics. In: S.F. Chipman, J.W. Segal & R. Glaser (Eds.), *Thinking and learning skills. Vol 2: Current research and open questions*. Hillsdale: Lawrence Erlbaum, 1985, 127-139.

Heller, J. & F. Reif, Prescribing effective human problem solving processes: Problem description in physics. *Cognition and Instruction*, 1984, *1*, 191-203.

Hewson, P.W., Diagnosis and remediation of an alternative conception of velocity using a microcomputer program. *American Journal of Physics*, 1985, *53*, 684-690.

Hewson, M.G. & P.W. Hewson, Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching*, 1983, *20*, 731743.

Hewson, P.W. & M.G. Hewson, The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science*, 1984, *13*, 1-13.

Hul, F. van 't, F. Kulik, P. Lijnse & P. Snijders Blok, *De trage inbreker. Een computerspel over krachten en beweging*. R.U. Utrecht, 1987.

Jong, T. de, *Kennis en het oplossen van vakinhoudelijke problemen* (proefschrift). Helmond: Wibro, 1986.

Jong, T. de & M.G.M. Ferguson-Hessler, *Voorwaarden voor het succesvol oplossen van problemen*. Technische Hogeschool, Eindhoven, 1982, rapport nr. 30.

Joshua, S. & J.J. Dupin, Taking into account student conceptions in instructional strategy: An example in physics. *Cognition and Instruction*, 1987, *4*, 117-135.

Labudde, P., F. Reif & L. Quinn, Facilitation of scientific concept learning by interpretation procedures and diagnosis. *International Journal of Science Education*, 1988, *10*, 81-98.

Lijnse, P.L., Schoolbeeld of straatbeeld: Over onderzoek naar begripsmoeilijkheden van leerlingen bij het leren van mechanica. In: Werkgroep Natuurkunde Didaktiek, *Zoeklicht op de mechanica. Verslag van de konferentie 'Woudschoten'*. Utrecht, 1982, 57-80.

Linn, M.C., K.D. Sloane & M.J. Clancy, Ideal and actual outcomes from precollege Pascal Instruction. *Journal of Research in Science Teaching*, 1987, *24*, 467-490.

Minstrell, J., Explaining the 'at rest' condition of an object. *The Physics Teacher*, 1982, *20*, 10-14.

McCloskey, M., Naive theories of motion. In: D. Getner & L.A. Stevens (Eds.), *Mental models*. Hillsdale: Lawrence Erlbaum, 1983, 299-323.

Nussbaum, J. & S. Novick, Alternative frameworks, conceptual conflict and accomodation: toward a principled teaching strategy. *Instructional Science*, 1982, *11*, 183-200.

Pfundt, H. & R. Duit, *Bibliography: Students' alternative frameworks and science education. Institute for Science Education*, Kiel, 1988. IPN Reports in brief nr. 34.

Posner, G.J., K.A. Strike, P.W. Hewson & W.A. Gertzog, Accomodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 1982, *66*, 211-227.

Rosenquist, M.L. & L.C. McDermott, A conceptual approach to teaching kinematics. *American Journal of Physics*, 1987, *55*, 407-415.

Reif, F., Acquiring an effective understanding of scientific concepts. In: L.T.H. West & A.L. Pines (Eds.), *Cognitive structure and conceptual change*. New York: Academic Press, 1985, 133-151.

Reif, F., Interpretation of scientific or mathematical concepts: Cognitive Issues and instructional implications. *Cognitive Science*, 1987, *11*, 395-416.

Reif, F., J.H. Larkin & G.C. Brackett, Teaching general learning and problem solving skills. *American Journal of Physics*, 1976, *44*, 212-217.

Schoenfeld, A.H. & D.J. Herrmann, Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 1982, *5*, 484-494.

Schoenfeld, A.H., *Mathematical problem solving*. New York: Academic Press, 1985.

Trowbridge, D.E. & L.C. McDermott, Investigation of student understanding of the concept of velocity in one dimension. *American Journal of Physics*, 1980, *48*, 1020-1028.

Trowbridge, D.E. & L.C. McDermott, Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics*, 1981, *49*, 242-253.

Vanderlocht, M., *Selectiekennis en het selectieproces bij het oplossen van natuurkundige problemen*. K.U. Leuven, Afdeling Didactiek, ISSAD-project, 1989, voorlopig manuscript.

Vanderlocht, M. & J. Van Damme, Probleemaanpak en het oplossen van natuurkundige problemen. Een literatuuroverzicht. *Pedagogische Studiën*, 1989a, *66*, 97-106.

Vanderlocht, M. & J. Van Damme, Fouten en knelpunten bij het oplossen van problemen over de projectielbeweging. Een systematische analyse van 152 schriftelijke oplossingen. *Tijdschrift voor Didactiek der β*-wetenschappen, 1989b, *7*, 115-137.

Vegting, P., Voorwetenschappelijke begrippen bij leerlingen wat betreft kracht. In: Werkgroep Natuurkunde Didaktiek, *Verslag van de GIREP-conferentie 1984: The many faces of teaching and learning mechanics*. Utrecht, 1985, 10-21.

White, B.Y., Designing computer games to help physics students understand Newton's laws of motion. *Cognition and Instruction*, 1984, *1*, 69-108.

Curriculum vitae

M. Vanderlocht (1962) studeerde psychologie, specialisatie schoolpsychologie en studie- en beroepsoriëntering, aan de K.U. Leuven. Sinds 1986 verricht hij in het kader van het ISSAD-project onderzoek naar het leren oplossen van natuurkundige problemen.

J. Van Damme (1946) is als hoogleraar verbonden aan de Afdeling Didactiek van de K.U. Leuven. Hij is er hoofd van het Onderzoekscentrum voor Secundair en Hoger Onderwijs.

Correspondentieadres: M. Vanderlocht en J. Van Damme, Pedagogisch Instituut, Vesaliusstraat 2, B-3000 Leuven.

Manuscript aanvaard 22-8-'89

Summary

Vanderlocht, M. & J. Van Damme. 'Knowledge and problem solving in physics,' *Pedagogische Studiën*, 1990, *67*, 70-81

Students' achievement in problem solving in physics is highly determined by the quality of their knowledge bases. In this article we treat four aspects that are important in teaching an adequate knowledge base in physics. These aspects are: operational knowledge, selection knowledge, differentiation between related concepts and intuitive preconceptions. For each topic we firstly explain the meaning and the relevance. Furthermore, we discuss the results from descriptive and prescriptive research.

Onderwijsontwikkelingen in de USSR

Tegen de achtergrond van de spectaculaire ontwikkelingen in het oostblok werd op vrijdag 10 november 1989 in de jaarbeurshal te Utrecht een seminar 'Onderwijsontwikkelingen in de USSR' gehouden. Het seminar was georganiseerd door T. van der Meer van de Algemene Hogeschool Amsterdam en B. Kleekamp van het Nascholingscoördinaat van de Universiteit van Amsterdam. Gastsprekers waren prof.dr. V.V. Davydov, dr. V.S. Sobkin, en dr. E.D. Dneprov, allen verbonden aan de Pedagogische Afdeling van de Academie der Pedagogische Wetenschappen in Moskou. De bedoeling van het seminar was Nederlandse docenten in te lichten omtrent de meest recente inhoudelijke onderwijsvernieuwingen in de USSR.

De spits werd afgebroken door Davydov, die in een omstandig betoog zijn voor de kenner inmiddels vertrouwde ideeën nogmaals uiteenzette. Dat wil zeggen, dat hij betoogde dat het huidige onderwijs te zeer zou zijn gebaseerd op aanschouwelijkheid en het leren van empirische generalisaties, waarbij de leerlingen aan het leren van werkelijk theoretisch denken niet zouden toekomen. Dit theoretische denken kan de leerling zich alleen eigen maken indien hij of zij inzicht krijgt in de wordingsgeschiedenis en fundamentele uitgangsbegrippen van een bepaalde leerstof. Als didactisch middel stelde hij voor de leerstof zo vroeg mogelijk te symboliseren met behulp van schema's en modellen, terwijl hij de effecten van zijn aanpak wilde evalueren met behulp van 'vormende experimenten'. De lezer die deze samenvatting enigszins vaag vindt, heeft gelijk, maar hoeft zich niet benadeeld te vinden ten opzichte van de toehoorders bij Davydovs lezing die gespeend was van verhelderende voorbeelden. Davydovs lezing was derhalve te weinig informatief voor de kenner van zijn werk en te vaag voor de oningewijde.

Opmerkelijk was wel dat hij nu voor het eerst zijn ideeën plaatste tegen de achtergrond van een nieuwe visie op leerling, onderwijs en maatschappij. Het onderwijs zou gehumaniseerd en gedemocratiseerd dienen te worden en het zou in eerste instantie een zorg moeten zijn van maatschappelijke groeperingen samengesteld uit ouders en leraren en vrij moeten zijn van hinderlijke staatsbemoeienis.

Davydovs betoog werd gevolgd door een lezing van de socioloog Sobkin, die verslag deed van een recentelijk door hem uitgevoerd enquête-onderzoek onder onderwijsgevenden, ouders en leerlingen. Sobkin overlaadde de toehoorders in de beste Russische traditie met een duizelingwekkende reeks van percentages, maar het beeld dat uit zijn cijferrijke en anderhalf uur durende verslag oprees was uiterst vertrouwd. Ook in de Sovjetunie constateren ouders een duidelijke achteruitgang in het niveau van het onderwijs en een schrijnend gebrek aan discipline in de school. Ook in de Sovjetunie kampen de leerkrachten met een gebrek aan waardering en een ondermaatse bezoldiging, terwijl zij de ouders van hun leerlingen een toenemende onverschilligheid verwijten. Ook in de Sovjetunie, ten slotte, zijn de leerlingen ontevreden over zowel de leerstof als de leraren.

Het is de vraag wat deze bevindingen betekenen. Sobkins cijfers weerspiegelden immers slechts de meningen van de door hem ondervraagde leerkrachten, ouders en leerlingen. Dat deze personen over het algemeen niet tevreden zijn en dat de oudere generatie een duidelijke achteruitgang constateert, lijkt een fenomeen van alle tijden. Pas tegen de achtergrond van een analyse van de daadwerkelijke situatie — bijvoorbeeld een regelmatige toetsing van het kennisniveau van de leerkrachten en leerlingen — krijgen dergelijke klachten enig reliëf. Helaas ontbrak dit element aan Sobkins lezing.

Ten slotte sprak de historicus Dneprov over de geschiedenis van het sovjetonderwijs. Zijn betoog was zowel door de presentatie als door de inhoud verreweg het smakelijkste gerecht van de dag. Na geconstateerd te hebben dat de zegswijze dat elk volk de regering krijgt die het verdient voor zijn medeburgers veel te

hardvochtig is, stak Dneprov van wal met een gepassioneerde analyse van verschillende fases uit de geschiedenis van het sovjetonderwijs en de sovjetpedagogiek. Na dit treurig stemmende overzicht kwam hij tot een schets van de nu levende radicale onderwijsidealen. Fel keerde Dneprov zich hierbij tegen elk partij- of staatsbemoeienis met het onderwijs en hij pleitte voor een vorm van economische zelfstandigheid voor de school. Daarnaast achtte hij het van cruciaal belang dat het onderwijs gedepolitiseerd en gehumaniseerd zou worden. In Dneprovs visie behoren schooluniformen, politieke leuzen in de klas, hagiografieën van politieke leiders, liedjes over vadertje Lenin, en andere vaste ingrediënten van het sovjetschoolleven binnenkort tot het verleden. Grimmig schetste hij dat de plannen hiertoe door partij-organen en ministeries gesaboteerd worden en zijn slotsom was dat hij niet optimistisch was over de mogelijkheden van een daadwerkelijke snelle vernieuwing van het sovjetonderwijs. Dneprovs vurige betoog vormde een interessant slot van een niet geheel geslaagde dag. Doel van de dag was ingelicht te worden over recente inhoudelijke onderwijsontwikkelingen in de USSR en om over deze vernieuwingen van gedachte te wisselen. Ook spraken de organisatoren de hoop uit dat over de gevolgen van deze koerswijzigingen voor de opleidingen voor het onderwijzend personeel een vruchtbare discussie gevoerd zou kunnen worden. Om verschillende redenen is het hier echter niet van gekomen.

Ten eerste verzuimde de voorzitter drs. Th. Liket de drie gastsprekers in toom te houden, waardoor zij allen hun spreektijd verdubbelden en er in het programma geen ruimte overbleef voor de geplande discussies na afloop van elke lezing. Dit onafgebroken en urenlang aanhoren van lezingen is uitermate slopend en het dalende aantal in gebruik zijnde koptelefoons (voor de simultaanvertaling) correspondeerde volgens ons dan ook niet met een evenredig toenemende beheersing van de Russische taal. Het was pas in de namiddag — na afloop van de drie lezingen en enige verouderde lichtbeelden van vrolijk zingende Russische leerlingen — dat er gelegenheid was om gedurende een uur vragen te stellen. Ook hier kwam het echter niet tot een werkelijke uitwisseling van gedachten, aangezien de gasten naar oud Russisch gebruik elke vraag aangrepen voor het lanceren van een lang betoog. Vaak ook meende een tweede gastspreker het zojuist door de eerste gastspreker gegeven antwoord te moeten aanvullen, waarna de derde gastspreker nog een kleine nuancering aanbracht. Van de beoogde gedachtenwisseling was dus niet echt sprake.

Ten tweede zijn de beoogde of gerealiseerde inhoudelijke onderwijsvernieuwingen eigenlijk maar summier aan bod gekomen in de drie lezingen. Met name de laatste twee sprekers deden veeleer verslag van de nu heersende verwarring in het sovjetonderwijs en van de taaie politieke strijd die gaande is om veranderingen te kunnen bewerkstelligen. Zodoende was er te weinig materiaal voor een inhoudelijke discussie en kwamen de gevolgen van beoogde koerswijzigingen voor de opleidingen voor onderwijzend personeel in het geheel niet aan bod. In de loop van de dag werd de doelstelling te berichten over recente inhoudelijke vernieuwingen steeds meer uit het oog verloren en kwam het accent te liggen op de strijd die onze sovjetcollega's op dit moment leveren om een aantal vrijheden te verwerven die hier reeds lang gemeengoed zijn. Deze accentverschuiving was natuurlijk heel begrijpelijk in het licht van de huidige politieke context in de USSR, maar tevens betreurenswaardig gezien het doel van de bijeenkomst.

R. van der Veer
(Rijksuniversiteit Leiden)

Mededelingen

Stichting Langeveld-lezing

De Stichting Langeveld-lezing stelt zich ten doel de herinnering aan Professor dr. M.J. Langeveld levend te houden door twee-jaarlijks een lezing te organiseren op één van de terreinen waarop hij in zijn wetenschappelijke loopbaan werkzaam is geweest.

Op 26 april 1990 om 16.00 uur wordt de eerste Langeveld-lezing gehouden in de Pieterskerk in Utrecht door Prof.dr. M. van Manen, University of Alberta, Edmonton (Canada), onder de titel "In Loco Parentis".

Aanmeldingen en inlichtingen bij mevrouw L. van Toledo, Interdisciplinair Onderzoeksinstituut van de Rijksuniversiteit Utrecht, Heidelberglaan 1, Utrecht, Tel.: 030-534834.

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
14e jaargang, nr. 6, 1989

De mate van implementatie als evaluatiecriterium, door H.A.M. Franssen en T.W.H.M. Stoverinck
Tussen retoriek en dagelijkse praktijk: de micropolitiek van de school, door P.J.P. van den Bos en C.J. de Brabander
De stagiair als student; over leerhoudingen in de schoolstage, door J.A. de Jong

Tijdschrift voor Onderwijsresearch
14e jaargang, nr. 5, 1989

Schrijfvaardigheid en het schrijf-denkproces, door I. Breetvelt
Effecten van een metacognitief trainingsprogramma, door H.A.J. Biemans
Vergelijking van scoringsprocedures bij schaalconstructie geïllustreerd met de vragenlijst van Slade en Jenner, door P.M. Kroonenberg, B.C.S. Pangemanann en D. Bakker
Notities en Commentaren
De vorming van de controlegroep bij een intact veldexperiment, door E.M. Scholte
Programmering van onderwijsonderzoek: een kritisch commentaar op het beleid van SVO, door W.Th.J.G. Hoeben

Tijdschrift voor Orthopedagogiek
28e jaargang, nr. 12, 1989

In memoriam prof.dr. I.C. van Houte, door A.J. Wilmink
Kinderen met leer- en opvoedingsmoeilijkheden, door R. de Groot
Vertekende beeldvorming na 40 jaar lom-onderwijs, door A.J. Wilmink
Enkele ervaringen rond het lom-onderwijs uit de periode 1960-1972, door H. Menkveld
Dagboek van een schoolleider, door J.J. van der Sluis
Een kwart eeuw bezorgdheid over groei van het lom-onderwijs, door K. Doornbos

Ontvangen boeken

Leseman, P.P.M., *Structurele en pedagogische determinanten van schoolloopbanen* (dissertatie). Rotterdamse Schooladviesdienst, Project Onderwijs en Sociaal Milieu, Rotterdam, 1989.

Roede, E., *Explaining student investment* (dissertatie). SCO/U.v.A., Amsterdam, 1989, *f* 25,—.

Verheij, F., H. van Loon m.m.w. J.E. de Boer, *Intensieve residentiële behandeling van kinderen: een kinderpsychiatrische kliniek ten voorbeeld*. Van Gorcum, Assen, 1989, *f* 49,50.

Vincent, H.R., *Integraal denken. Vervreemding en de humanisering van arbeid, onderwijs en politiek*. Acco, Amersfoort, 1989, *f* 29,-.

Vleuten, C.P.M. van der, *Naar een rationeel systeem voor toetsing van studieprestaties in probleemgestuurd medisch onderwijs* (dissertatie). Uitgeverij Thesis, Amsterdam, 1989, *f* 43,50.

Vogel, P. (red.), *De school, keuzen en kansen*. Coutinho, Muiderberg, 1989, *f* 29,50.

Wiegers, A.M. & F.A. Swets-Gronert (red.), *Samen hulpverlenen bij kindermishandeling*. Een overzicht van de samenwerkingsverbanden. Van Loghum Slaterus, Deventer, 1989.

Wit, J. de, G. van der Veer, *Psychologie van de adolescentie*. Ontwikkeling en hulpverlening (vijftiende, herziene druk). Intro, Nijkerk, 1989, *f* 44,50.

Evolutie in onderwijsevaluatie

Ten geleide

In de tweede helft van de jaren zeventig vond in Nederland een viertal, door de overheid geïnitieerde onderwijsinnovatieprogramma's plaats: het middenschool experiment, het open school proefproject, het innovatieprogramma basisonderwijs en projecten op het terrein van het zogenoemde participatieonderwijs. Ten aanzien van de drie eerstgenoemde innovatieprogramma's werden activiteiten ontplooid die met evaluatie van doen hadden. Er verschenen nota's die het woord evaluatie in de titel hadden, er werden onderzoeksprogramma's opgesteld waarin – verschillend per innovatieprogramma – een minder of meer nadrukkelijk evaluatieperspectief aanwezig was en er werd een groot aantal, voor het merendeel kleinschalige, onderzoeksprojecten uitgevoerd.

Wanneer men achteraf de balans opmaakt, hetgeen onder meer gebeurd is in het kader van meta-evaluatieve studies (Scheerens, 1983a en b; Hoeben, 1985), moet geconcludeerd worden dat er van een systematische, wetenschappelijk verantwoorde evaluatie van de desbetreffende innovatieprogramma's weinig terecht is gekomen. Wetenschappers oordeelden zeer kritisch over de kwaliteit van het uitgevoerde onderzoek en beleidsvoerders en practici vonden het onderzoek weinig relevant. Het over het algemeen weinig succesvolle karakter van deze pogingen tot evaluatie kon gedeeltelijk worden toegeschreven aan een gebrek aan ambachtelijkheid bij de uitvoerders en aan de moeilijke grijpbaarheid van het evaluatie-object.

Tevens werd aangetekend dat de organisatorische uitvoeringscondities verre van optimaal waren. Dit bleek uit:

- onduidelijke beslissingsstructuren rondom de programmering van het onderzoek, hetgeen leidde tot langdurige overlegprocedures tussen departementsambtenaren, semi-autonome programma-adviescommissies, de SVO, innovatiecommissies, en incidenteel ook onderzoeksinstituten;
- een daarmee samenhangend verantwoordelijkheidsvacuüm; binnen de veelheid aan instanties dat betrokken was bij de opzet en uitvoering van het onderzoek was er geen die een reputatie te winnen of te verliezen had bij dit soort onderzoek;
- versnippering en een gebrek aan daadkracht bij de uitvoering van de onderzoeksprogramma's;
- grote weerstanden bij scholen om aan het evaluatie-onderzoek mee te werken.

Daarbij kwam dat de 'politieke wil' om 'hard', d.w.z. output-gericht en wetenschappelijk evaluatie-onderzoek te doen uitvoeren, bij belangrijke betrokkenen, zoals departementsambtenaren en leden van onderwijsinnovatiecommissies in de periode in kwestie afwezig was. De bedoeling van het onderzoek werd geschetst als 'het opleveren van informatie'. Onderzoekers die het waagden om aan de door hen verzamelde gegevens evaluatieve conclusies te verbinden, werden op de vingers getikt. In beleidsnotities van de minister werd gerept over de schadelijkheid voor het innovatieproces van vooraf opgestelde evaluatiecriteria en innovatiecommissies (in het bijzonder de innovatiecommissie middenschool) waarschuwden voor evaluatie waar het innovatieproces nog niet aan toe was, voor de onvoldoende tot wasdom gekomen evaluatiemethodologie evenals het ontbreken van voldoende onderwijskundige kennis.

In het begin van de jaren tachtig zette in beleidskringen een kentering in over het denken over onderwijsevaluatie. Over onderwijskwaliteit, effectiviteit van vernieuwingsprogramma's en rendement mocht niet alleen opeens wèl hardop gesproken worden, het werden zelfs de centrale issues. Verschijningsvormen van deze geleidelijke verandering waren de nota Kwaliteit van het Onderwijs (1981), het ontwerpprogramma voor de evaluatie van het Onderwijsvoorrangsbeleid van Hofstee, Meijnen en Koster (1984), het ten departemente opgestelde evaluatieplan voortgezet onderwijs (1984), de nota gedrags- en maatschappijwetenschappen (1984) waarin programma-evaluatie als één van de stimuleringsgebieden voor onderzoek werd genoemd, de geleidelijke start van het periodieke peilings-

onderzoek (dat sinds 1980 werd voorbereid door beleidsplanners van O&W, de SVO en het CITO) en de oprichting door de Minister van O&W van de klankbordcommissie Programma Evaluatie, in 1985.

Ook de SVO profileerde zich in deze periode nadrukkelijker als coördinator van onderwijsevaluatie, onder meer door een afzonderlijke bureau-afdeling voor dit werkterrein in het leven te roepen.

Als vermoedelijke oorzaken van de mentaliteitsverandering op het terrein van onderwijsevaluatie zijn (in volgorde van belangrijkheid) te noemen: het verhoogde doelmatigheidsbewustzijn in het overheidsbeleid (bezuinigingen, deregulering, output-financiering e.d.); contacten van invloedrijke beslissers (c.q. departementsambtenaren) met buitenlandse evaluatie- en assessment-programma's; ontwikkelingen in de evaluatiemethodologie en -technologie en de aanwezigheid van wetenschappers (te beginnen met A.D. de Groot) die altijd al in waren voor output-georiënteerd en wetenschappelijk evaluatie- en assessmentonderzoek – andere namen die in dit verband genoemd moeten worden zijn Hofstee, Creemers, Meijnen, Hoeben, Scheerens en Koster.

In dit gewijzigde klimaat ten aanzien van onderwijsevaluatie werden in dezelfde periode (eerste helft van de jaren tachtig) nieuwe landelijke onderwijsinnovatieprogramma's in de steigers gezet, te weten het onderwijsvoorrangsbeleid, vernieuwing van de eerste fase van het voortgezet onderwijs en het informatica-stimuleringsplan. Bij elk van deze programma's zijn evaluaties in voorbereiding of reeds in uitvoering.

In dit themanummer wordt verslag gedaan van de voorbereiding en uitvoering van de eerste fasen van deze nieuwe generatie onderwijsevaluaties. Er zijn dus bijdragen over de voorgenomen evaluatie van de eerste fase van het voortgezet onderwijs, het onderwijsvoorrangsbeleid en het (onderwijsdeel van het) informatica-stimuleringsplan. Daarnaast is er een bijdrage over het periodieke peilingsonderzoek.

Scheerens beschrijft in zijn bijdrage de context, de methodologie en de resultaten van twee evaluatiestudies die zijn uitgevoerd in het kader van het Informatica Stimuleringsplan (INSP): een evalueerbaarheidsstudie en een sobere toekomstgerichte evaluatie van het INSP. Aan het eind van zijn artikel gaat hij, ondanks een aantal vernieuwende elementen in de aanpak, in op een aantal kritische kanttekeningen die te plaatsen zijn bij een evaluatie die meer weg heeft van een onderbouwd advies dan van een traditionele empirische benadering.

Nijhof beschrijft in de tweede aansluitende bijdrage een deelevaluatie van het INSP. Eén van de doelen van het INSP was het stimuleren van informatiekunde in het beroepsgerichte onderwijs o.a. door aan te sluiten bij nieuwe de ontwikkelingen in beroepen en functies. Een drietal deelprogramma's is geëvalueerd, daarnaast is gebruik gemaakt van informatie van externe (bedrijfs-)opleidingen. De conclusie is dat de primaire doelen niet ten volle uit de verf zijn gekomen, maar dat belangrijke aanzetten en stimulansen zijn gegeven voor de ontwikkeling van informatiekunde en de aansluiting op de arbeidsmarkt.

Peschar schetst de overwegingen en uitgangspunten voor een evaluatie van de basisvorming, waarbij gekozen is voor een loopbaanonderzoek in cohorten met drie meetpunten. Hoewel de discussie over de basisvorming voortduurt lijkt er in evaluatief opzicht winst geboekt te zijn: een doordacht grootschalig vergelijkend onderzoeksprogramma naar de effecten van schoolkenmerken in het voortgezet onderwijs.

Onderwijsvoorrangsbeleid is gericht op een controversieel vraagstuk. Evaluatie van onderwijsvoorrangsbeleid (OVB) is zo mogelijk nog controversiëler. Van der Werf en Tesser beschrijven de context van het OVB, het onderzoekdesign, de discussie over de effectmaten en het uitvoeringsplan van de evaluatie. Uit een vooronderzoek worden enkele resultaten gepresenteerd. Opmerkelijk is dat de implementatie van het OVB op scholen vaak niet duidelijk is. Over de effecten hoeven we dan vooralsnog geen illusies te hebben.

De laatste bijdrage in de reeks heeft betrekking op de Periodieke Peiling van het Onderwijsniveau. Wijnstra schetst de werkwijze van het periodieke peilingsonderzoek (PPON) aan de hand van de uitgevoerde rekenpeiling in 1987. Peilingsonderzoek (ook wel assessment genoemd) is een bepaalde vorm van evaluatie-onderzoek die in een aantal opzichten duidelijk afwijkt van de voorgaande evaluatie-onderzoeken: het spoort geen directe

effecten van programma's op, maar probeert vast te stellen wat er gekend en gekund wordt door een populatie leerlingen in het licht van wat de samenleving wenselijk acht.

Nijhof en Scheerens sluiten de serie kort af met een aantal constateringen en observaties. Zo stellen zij vast dat evaluatieprogramma's meer produktgericht worden en dat het belang van evaluatie in grootschalige programma's bij overheid en politiek meer wordt onderkend en erkend. Daarnaast blijft uiteraard de vraag bestaan welk design het meest geschikt is. Vragen blijven er wel ten aanzien van het gebruik en de bruikbaarheid van de evaluatiegegevens. Sommige onderwijspolitieke doelen zijn niet grijpbaar voor welke vorm van evaluatie dan ook: sommige designs sterven in schoonheid.

Namens de redactie
W.J. Nijhof
J. Scheerens

Literatuur

Hoeben, W.Th.J.G., Evaluatie-onderzoek van het onderwijs. Stand van zaken. In: R. Halkes & R.G.M. Wolbert (red.), *Docent en methode*. Bijdrage tot de Onderwijsresearchdagen 1984 te Tilburg Lisse: Swets & Zeitlinger, 1985. (pp. 23-41)

Scheerens, J., *Evaluatie-onderzoek en beleid*. Methodologische en organisatorische aspecten. Harlingen: Flevodruk, 1983a.

Scheerens, J., *Het sectoronderzoek: onderwijsonderzoek in de marge van wetenschap en beleid?* Een studie naar het functioneren van het beleidsgerichte onderzoek in het kader van enkele landelijke onderwijs-innovatieprojecten. Harlingen: Flevodruk, 1983b.

De evaluatie van het onderwijsdeel van het Informatica Stimuleringsplan (INSP)

J. SCHEERENS
Vakgroep Onderwijsorganisatie en -management, Universiteit Twente

Samenvatting

In dit artikel worden de inhoud, methodologie en resultaten van twee evaluatiestudies van het Nederlandse Informatica Stimuleringsplan (INSP) besproken; het onderwijsdeel van het programma vormt daarbij het belangrijkste aandachtspunt. De eerste evaluatiestudie die wordt besproken is een evalueerbaarheidsanalyse, de tweede is een 'bescheiden toekomstgerichte evaluatie'. Elke studie werd uitgevoerd in ruwweg zes maanden tijd. De sterke en zwakke punten van de evaluatiebenadering die gebruikt werd bij de tweede studie worden besproken. Geconcludeerd kan worden dat het proactieve, toekomstgerichte en deels analytische (zowel als empirische) karakter van deze studie gerechtvaardigd was, gezien de ontwikkelingsfase van het te evalueren programma en de randvoorwaarden voor de evaluatie.

De opvallendste bevindingen zijn de erkenning van het belang van onderwijs en training voor alle onderdelen van Informatica Stimulering in de maatschappij en de onderkenning van de ernst van het probleem hoe grootschalige Informatica Stimulering in het onderwijs georganiseerd en bestuurd moet worden.

1 *Inleiding*

In dit artikel worden twee fasen van de evaluatie van het Informatica Stimuleringsplan (hier verder afgekort tot INSP) besproken. De voorliggende presentatie heeft vooral betrekking op het onderwijsdeel van de evaluatie. Aangezien de evaluatie van het onderwijsdeel een integraal onderdeel vormde van de evaluatie van het gehele INSP, komt ook het totaalontwerp van de INSP-evaluatie ruimschoots aan de orde.

De eerste fase van het INSP-evaluatie was een zogenoemde 'evalueerbaarheidsanalyse', waarbij moest worden vastgesteld in hoeverre en op welke wijze het INSP te evalueren was.

De tweede fase is aan te duiden als een beperkte, proactieve en formatieve evaluatie. Met het predicaat 'proactief' wordt tot uitdrukking gebracht dat de in het INSP op gang gebrachte ontwikkelingen werden getoetst aan prognoses over de toekomstige maatschappelijke betekenis van informatietechnologie. De beide fasen van de INSP-evaluatie werden uitgevoerd door commissies, waarvan de auteur deel uit maakte[1]. Het rapport van de evalueerbaarheidsanalyse verscheen in juni 1987, het rapport van de beperkte formatieve evaluatie in juni 1988.

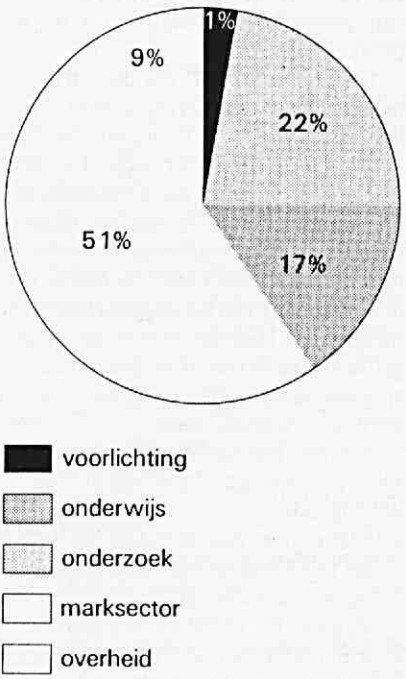
Eerst zal een korte typering worden gegeven van het INSP. Daarna worden de opzet en resultaten van de beide evaluaties besproken. In de discussie wordt niet alleen gereflecteerd op de inhoudelijke onderzoeksuitkomsten, maar ook op bijzonderheden van de toegepaste evaluatiemethodologie en de organisatie van beide projecten.

2 *Het Informatica Stimuleringsplan (INSP)*

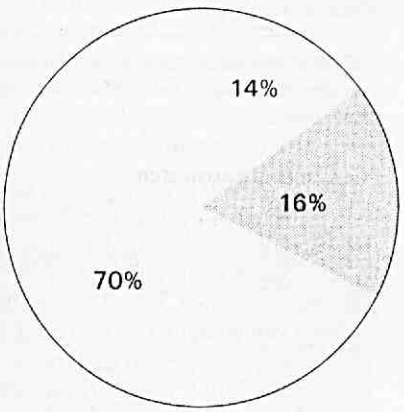
Het INSP ging van start in 1984. Aan het project werd meegewerkt door drie departementen: Economische Zaken, Onderwijs en Wetenschappen en Landbouw en Visserij. De algemene doelstelling van het INSP was het stimuleren van de maatschappelijke toepassing van informatietechnologie en het wegwerken van een veronderstelde achterstand op andere geïndustrialiseerde landen. Het project is formeel beëindigd in december 1988. Het totale budget bedroeg 1,7 miljard gulden. In Figuur 1 wordt een indruk gegeven van de verdeling van het totale budget over specifieke toepassingsgebieden en van de verdeling over 'sponsors'.

	EZ	O&W	LaVi	Totaal
voor-lichting	14			14
onder-wijs	140	190	50	380
onder-zoek	115	85	90	290
markt-sector	820		39	859
over-heid	108		49	157
totaal	1197	275	228	1700

Vooral stimulering marktsector



Ruim 70% gefinancierd door EZ



Bron: Ministerie van EZ, O&W, LaVi

Figuur 1 *Financiering INSP door deelnemende departementen per sector (in f miljoen)*

De vijf deelprogramma's van het INSP waren voorlichting (1% van het totale budget), onderwijs (22%), research (17%), stimulering van de marktsector (51%) en de stimulering van computertoepassingen bij de overheid (9%).

Uit Figuur 1 blijkt dat de marktsector het belangrijkste stimuleringsgebied was en dat het Ministerie van Economische Zaken de voornaamste financier was. Uit de figuur wordt ook duidelijk dat Economische Zaken niet minder dan 37% van het totale budget voor onderwijsstimulering voor zijn rekening nam (bijna de helft hiervan betrof overigens de stimulering van het zogenoemde irreguliere onderwijs). Hiermee wordt het hoge belang geïllustreerd dat werd toegekend aan stimulering van de aanbodkant van de economie in de vorm van 'human capital'.

Andere belangrijke kenmerken van het INSP, die niet zichtbaar zijn in Figuur 1, zijn de volgende:

- Niet minder dan 45% van het budget voor het INSP bestond uit reeds bestaande fondsen en subsidieregelingen die nu onder de vlag van het INSP werden gepresenteerd.
- Bijna 60% van de INSP-middelen werden besteed aan stimulering van de aanbodkant van informaticatoepassingen (met name aan de ontwikkeling en distributie van hard- en software).
- Elk van de 5 deelprojecten van het INSP was onderverdeeld in talrijke projecten en deelprojecten. Als geheel vormden deze deelprojecten een vrij losse structuur die op onderdelen wel redelijk, maar over het geheel genomen maar zwak gecoördineerd werd.

In het ontwerp voor het INSP worden twee hoofddoelstellingen van het *onderwijsgedeelte* genoemd:

- het vertrouwd maken van 'de burgers' met de informatietechnologie (o.a. burgerinformatica);
- het creëren van 'human capital' ten behoeve van de versterking van de marktsector, alsmede ten behoeve van de versterking van het beter en doelmatiger functioneren van maatschappelijke voorzieningen (p. 20). Verder werd in dit ontwerp ook 'een toename van het gebruik van de computer in het onderwijs' als (proces)doelstelling genoemd.

De structuur van het onderwijsgedeelte van het INSP laat een opbouw zien van enerzijds deelprogramma's die specifiek zijn voor een bepaalde onderwijscategorie (basisonderwijs, voortgezet onderwijs, lager en middelbaar beroepsonderwijs en hoger beroepsonderwijs) en anderzijds ondersteunende deelprogramma's. De ondersteunende deelprogramma's hadden betrekking op het creëren van een infrastructuur voor de ontwikkeling van educatieve software, scholing en nascholing van leerkrachten en onderwijsresearch op het terrein van toepassingen van informatietechnologie in het onderwijs. In Figuur 2 worden de budgetten voor elk van de subprogramma's van het onderwijsdeel van het INSP weergegeven.

Basis- en speciaal onderwijs	15
Voortgezet onderwijs, fase 1	25
Voortgezet onderwijs, fase 2	10
Lager en middelbaar beroepsonderwijs	75
Hoger beroepsonderwijs	43
Agrarisch onderwijs	56
Volwasseneneducatie	43
Infrastructuur educatieve software-ontwikkeling	40
Scholing en bijscholing van leerkrachten	53
Onderwijsonderzoek	10

Figuur 2 *Budgetten voor subprogramma's van het onderwijsdeel van het INSP over de periode 1984-1988*

Figuur 2 geeft een indruk van de verdeling van prioriteiten over de verschillende onderwijscategorieën en de ondersteunende activiteiten. Daarbij valt onder meer de betrekkelijk lage prioriteit op, die gegeven is aan vormen van computerondersteund onderwijs in het basis- en speciaal onderwijs. Wat niet uit Figuur 2 blijkt, maar wèl van belang is bij het verkrijgen van een juist beeld over de uitgaven binnen het onderwijsdeel van het INSP, is het feit dat 80% van het totale budget voor het onderwijsdeel besteed werd aan hardware (50%) en software (30%).

Evenals het gehele INSP was ook het onderwijsdeel onderverdeeld in honderden deelprojecten. Voor de coördinatie van al deze projecten was er een managementstructuur in het leven geroepen, die meer geprononceerd was dan bij andere hoofdonderdelen van het INSP. Deze bestond uit een projectgroep op

het Ministerie van Onderwijs en Wetenschappen en externe 'clustermanagers'. Uit de voortgangsverslagen over het INSP blijkt (onder meer uit de begrotingsoverzichten) dat het veel moeite kostte om al deze deelprojecten tijdig tot uitvoering te brengen. Voor verdere bijzonderheden over het onderwijsdeel van het INSP zij verwezen naar Plomp, Van Deursen en Moonen (1987).

De complexiteit, de beperkte tijdsduur die beschikbaar was voor implementatie (5 jaar) en het feit dat deelprojecten een zodanig losse structuur hadden dat het vaak moeilijk was om de bijdrage van deelprojecten aan de doelstellingen van het programma als geheel te relateren, zijn even zovele factoren die het onderwijsdeel van het INSP tot een moeilijk evaluatie-object maakten.

3 *Evalueerbaarheidsanalyse*

3.1 *Opzet en methoden*

In de evaluatieliteratuur heeft het idee van evalueerbaarheidsanalyse nog betrekkelijk weinig aandacht gekregen. Door sommige auteurs is evalueerbaarheidsanalyse als voorstadium voor grootschalig evaluatie-onderzoek voorgesteld ter bevordering van een doelmatige inzet van evaluatie-onderzoek (dat wil zeggen evaluatie gericht op de elementaire doel-middel relaties van het te evalueren programma en het vermijden van al te grote organisatorische risico's bij de vormgeving van de evaluatie) – vgl. Jung en Schubert, 1983; Scheerens, 1985. Bij enkele recente grootschalige Nederlandse evaluatieprojecten is evalueerbaarheidsanalyse gekoppeld aan een analytische evaluatie, welke voorafging aan het empirische evaluatie-onderzoek (vgl. Beem, Pijl & Scheerens, 1985; Beem 1988; Hofstee, Schaveling & Koster, 1987). Bij analytische evaluatie staat een rationele reconstructie van het te evalueren programma centraal, waarbij tevens een waarde-oordeel wordt uitgesproken over de causale assumpties die aan het programma ten grondslag liggen. Hoogerwerf (1984) spreekt in deze van de explicitering van de beleidstheorie. Men zou kunnen zeggen dat het verband tussen evalueerbaarheidsanalyse en analytische evaluatie wordt bepaald door de gemeenschappelijke noodzaak om het te evalueren programma te reconstrueren in termen van de elementaire middel-doel relaties. Evalueerbaarheidsanalyse gebruikt deze reconstructie voor het doen van praktische voorstellen inzake de uit te voeren empirische evaluatie. Analytische evaluatie gebruikt de reconstructie als basis voor het doen van waarderende uitspraken.

De gedachte om de evaluatie van het INSP vooraf te laten gaan door een evalueerbaarheidsanalyse werd gelanceerd door de Commissie Programma Evaluatie (een adviescommissie van het Directoraat Generaal Wetenschapsbeleid van het Ministerie van O&W). Deze commissie formeerde een (externe) werkgroep om de evalueerbaarheidsanalyse uit te voeren. De evalueerbaarheidsanalyse moest antwoord geven op de volgende vragen:

- in hoeverre is een overkoepelende evaluatie (d.w.z. het gehele programma omvattende evaluatie) haalbaar?
- op welke wijze zou de evaluatie uitgevoerd moeten worden?
- welke deskundigheid zou voor de evaluatie ingezet moeten worden en hoe zou de evaluatie georganiseerd moeten worden?

Bij de aanpak van de evalueerbaarheidsanalyse gebruikte de werkgroep een conceptueel kader dat bestond uit een onderscheid in '*ambitieniveaus*' bij programma-evaluatie, een onderverdeling van het INSP in verschillende aggregatieniveaus, een onderscheid in *evaluatieperspectieven* en een check-list om de evalueerbaarheid van deelprogramma's te bepalen.

Bij het onderscheid in ambitieniveaus stelde de werkgroep vast dat programmadoelstellingen het ankerpunt zijn voor evaluatie. Gegeven deze programmadoelstellingen dient evaluatie vast te stellen:

- of de doelstellingen bereikt zijn;
- in hoeverre doelrealisatie inderdaad ondubbelzinnig aan het programma kan worden toegeschreven;
- of de resultaten ook op de meest efficiënte wijze bereikt zijn.

Deze opsomming van vragen is te beschouwen als een qua ambitieniveau opklimmende reeks, waarbij de laatstgenoemde vragen veronderstellen dat eerder genoemde vragen beantwoord zijn. Evalueerbaarheidsanalyse zou kunnen uitwijzen dat bijvoorbeeld wel de eerste vraag te beantwoorden is, maar dat het tevens beantwoorden van de tweede en derde vraag te moeilijk of te duur zou zijn. Overi-

gens werd door de werkgroep onder ogen gezien dat in sommige gevallen de effectvraag (eerste vraag) voorlopig niet te beantwoorden zou zijn en men in feite niet veel meer zou kunnen doen dan nagaan of het programma of deelprogramma adequaat ingevoerd werd. Controle op implementatie werd door de werkgroep echter meer gezien als voorbereiding op programma-evaluatie, dan als programma-evaluatie op zich.

In het conceptuele kader dat de werkgroep gebruikte voor de evalueerbaarheidsanalyse werd verder het INSP uiteengelegd in een viertal aggregatieniveaus:
- niveau 1, het INSP als geheel;
- niveau 2, de vijf hoofdonderdelen van het INSP (voorlichting, onderwijs, research, stimulering marktsector en informatietechnologie-toepassingen bij de overheid);
- niveau 3, zo'n 30 deelprogramma's verdeeld over de vijf hoofdgebieden;
- niveau 4, het projectniveau.

De werkgroep koos niveau 3 als het meest geschikte niveau voor de evalueerbaarheidsanalyse, omdat de eerste twee niveaus te algemeen en het vierde niveau te specifiek werden gevonden. Verder onderscheidde de werkgroep drie evaluatieperspectieven (met evaluatieperspectief wordt hier bedoeld een zienswijze op evalueren die leidt tot een bepaalde algemene benadering):
- het *beleidsimmanente perspectief*, waarbij gekapitaliseerd wordt op de officiële doelstellingen en werkwijzen van het programma;
- het *beleidsanalytische perspectief*, waarbij ook de uitgangspunten en de logische samenhang van het programma ter discussie worden gesteld;
- het *meta-evaluatieve perspectief*, waarbij evaluatie het karakter heeft van een beoordelende synthese van reeds uitgevoerde deelevaluaties.

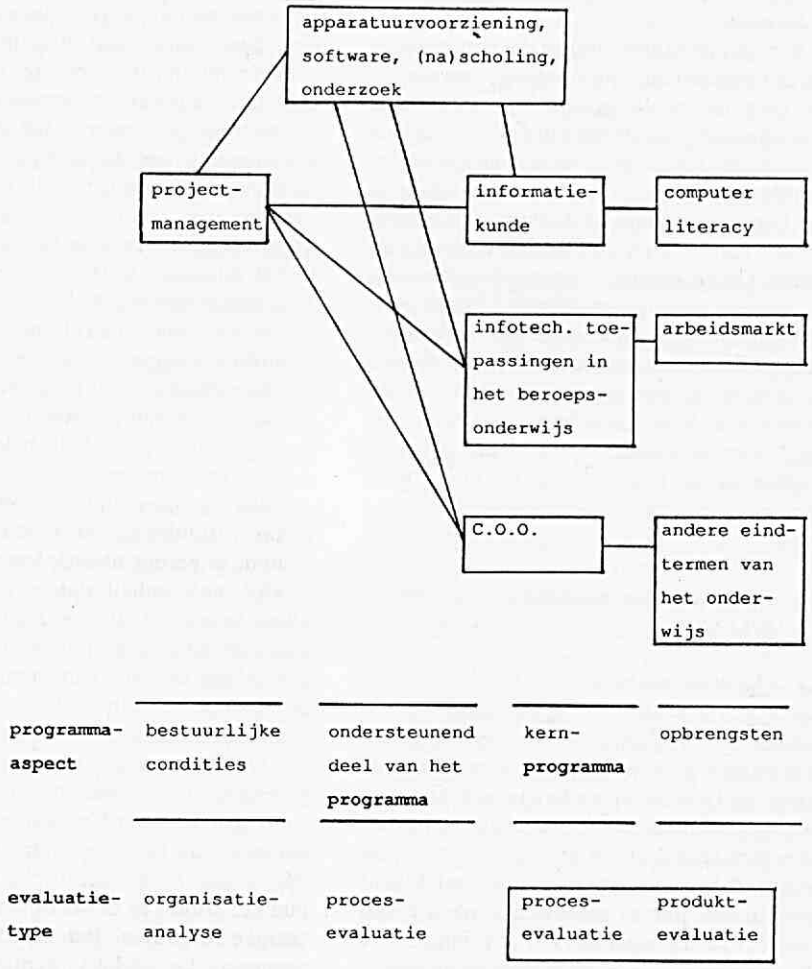
Deze perspectieven werden door de werkgroep gebruikt als alternatieve mogelijkheden bij de evaluatie van deelprogramma's.

Het laatste onderdeel van het conceptuele kader dat gebruikt werd bij de evalueerbaarheidsanalyse was een check-list om de evalueerbaarheid van deelprogramma's te bepalen. De hoofdcategorieën van deze check-list waren als volgt:
- beschrijving van het deelprogramma in termen van doelen en hoofduitvoeringswegen;
- formulering van het evaluatieprobleem; keuze van het meest in aanmerking komende evaluatieperspectief;
- evalueerbaarheid in de zin van de drie kernvragen bij programma-evaluatie (toetsbaarheid van effecten, toeschrijving aan het programma en efficiency-overwegingen);
- evalueerbaarheid in de zin van 'voorbereidingen op programma-evaluatie', zoals controle op budgetbesteding en invoering;
- algemene opzet voor de evaluatie van elk deelprogramma, inclusief een indicatie van de benodigde expertise, financiële middelen en uitvoeringstijd.

3.2 *Resultaten*

Zoals eerder aangegeven werd de evalueerbaarheidsanalyse uitgevoerd op het niveau van de ongeveer 30 deelprogramma's van het INSP. Hier zullen alleen de resultaten voor het hoofdgebied onderwijs worden samengevat. De 10 deelprogramma's binnen het hoofdgebied onderwijs zijn al eerder vermeld en samengevat in Figuur 2. In Figuur 3 worden de voornaamste aspecten van de onderwijsprogramma's weergegeven, nl. organisatorische condities, de ondersteunende sub-programma's (infrastructuur programmatuurontwikkeling, scholing en onderwijsresearch), het kernprogramma verdeeld over de verschillende onderwijssoorten en de beoogde resultaten. Per aspect is in Figuur 3 eveneens het meest in aanmerking komende type evaluatie-onderzoek weergegeven.

De conclusies van de evalueerbaarheidsanalyse kwamen er in hoofdlijnen op neer dat een organisatie-analyse van de project-managementstructuur relevant en haalbaar werd geacht, dat verhoopte opbrengsten, zoals een vergroting van 'computer literacy' in het voortgezet onderwijs en arbeidsmarkteffecten van informatietechnologie-toepassingen in het beroepsonderwijs in principe meetbaar werden geacht, maar op een later tijdstip dan de einddatum van het INSP en dat procesevaluatie van de ondersteunende deelprogramma's haalbaar en wenselijk was. Ten slotte werd geconcludeerd dat het zeer moeilijk zou zijn om de bijdrage van computer ondersteund onderwijs aan de algehele kwaliteit en doelmatigheid van het onderwijs vast te stellen.



Figuur 3 *Structuur van het onderwijsdeel van de INSP en evaluatietypen*

Verder werden er algemene evaluatieplannen opgesteld voor die onderdelen, welke evalueerbaar werden geacht. De voornaamste van deze voorstellen waren:

a. proces-evaluatie bedoeld om vast te stellen of de INSP-deelprogramma's uitgevoerd werden zoals gepland; waarbij de aanbeveling werd gedaan dat de meeste van deze 'implementatie-checks' door de onderwijsinspectie zouden worden uitgevoerd;
b. effect-evaluatie van het NIVO-project – een grootschalig project in het voortgezet onderwijs, gericht op verhoging van de computer literacy – door middel van een internationaal vergelijkend onderzoek, dat gekoppeld zou kunnen worden aan het onderzoek van de International Association for the Evaluation of Educational Achievement (IEA) naar computers in het onderwijs;
c. verschillende longitudinale onderzoeken gericht op het vaststellen van effecten op het beroepsmatig functioneren van (ex-) leerlingen uit het beroepsonderwijs, speciaal in beroepsvelden waar informatietechnologie van cruciale betekenis is geworden;
d. organisatie-analyse van de INSP-managementstructuur voor het hoofdgebied onderwijs; o.a. met het oog op het ontwerp

van toekomstige managementstructuren voor onderwijsinnovatie op het terrein in kwestie.

Eén van de aanbevelingen die voortvloeide uit de evalueerbaarheidsanalyse – en die betrekking had op de totale INSP-evaluatie en niet alleen op het onderwijsdeel – was het voorstel om niet direct over te gaan tot een integrale empirische evaluatie, maar om eerst een beperkte analyse uit te voeren, die mede gericht zou moeten zijn op het toekomstige beleid. Op die manier – zo was de redenering – zou een doelmatiger selectie kunnen worden gemaakt uit de vele mogelijke deelevaluaties, nl. door onderdelen te selecteren die van grote betekenis werden geacht voor het verdere beleid ter stimulering van de 'informatisering' van de maatschappij. Deze beperkte analyse is de tweede fase van de INSP-evaluatie die in dit artikel besproken wordt.

4 *Proactieve, formatieve evaluatie van het INSP*

4.1 *Opzet en methoden*

De beperkte analyse die in het kader van de evalueerbaarheidsanalyse als eerstvolgende stap bij de evaluatie van het INSP werd aanbevolen, had enkele bijzondere kenmerken. Deze kenmerken waren een proactieve en formatieve gerichtheid en een deels analytisch, deels empirisch karakter. De proactieve gerichtheid hield in dat, min of meer onafhankelijk van het INSP, werd nagegaan wat de belangrijkste terreinen waren voor informatietechnologiestimulering op langere termijn. Hieraan werd vormgegeven door het op ruime schaal interviewen van 'stake-holders' op het terrein van de informatietechnologie in het bedrijfsleven, de wetenschap en bij de overheid en door buitenlandse panels van experts te consulteren. Met de kennis die in deze oriënterende fase van de evaluatie werd verkregen, werd vervolgens een selectie gemaakt van zes thema's, die van strategisch belang werden geacht voor de toekomstige stimulering van informatietechnologie. Deze thema's waren: produktieautomatisering, bedrijfsoverschrijdende logistiek, software en de softwaresector, management van complexe informatiesystemen, courseware en informatietechnologie-stimulering in regulier en irregulier beroepsonderwijs op lager en middelbaar niveau. Voor elk van deze thema's werd een deelstudie uitgezet bij gespecialiseerde bureaus en instellingen voor wetenschappelijk onderzoek. Bij elk van deze deelstudies stond enerzijds opnieuw een verkenning van toekomstige ontwikkelingen centraal, maar werd anderzijds een evaluerende terugblik geworpen op het INSP. Naast de dominante proactieve gerichtheid was er bij deze evaluatie van het INSP dus ook sprake van een meer retroactieve evaluatie van hetgeen in het INSP tot stand was gebracht. Meer in het bijzonder waren de deelstudies op de volgende aspecten gericht:

- analyse van het betreffende onderwerp;
- globale waardering van de meest relevante onderdelen van het INSP-beleid;
- voor zover van toepassing, een schets van voorgestaan post-INSP-beleid;
- voor zover mogelijk, een vergelijking met ontwikkelingen in het buitenland;
- een formulering van oplossingsrichtingen voor de geconstateerde knelpunten.

Het analytische karakter van de evaluatie kwam vooral tot uitdrukking in het zwaarwegende en richtinggevende belang dat werd toegekend aan de visies van terzake kundigen op de toekomst van informatietechnologie. Daarnaast had deze beperkte evaluatie van het INSP een duidelijk empirisch karakter, met name op die onderdelen waar het INSP 'retrospectief' werd geëvalueerd. Overigens berustte ook het meer opiniërende deel van deze studie op systematische gegevensverzameling, vooral in de vorm van interviews (in totaal werden er meer dan 600 personen geïnterviewd). De algehele oriëntatie van deze beperkte evaluatie was formatief, dat wil zeggen dat de uitkomsten vooral waren bedoeld om onderbouwde suggesties aan te voeren voor het verdere I.T.-stimuleringsbeleid en minder om finale waardeoordelen uit te spreken over het INSP. De tijdsplanning van de evaluatie was hierop ook afgestemd: het eindrapport diende te worden opgeleverd op een tijdstip dat nog enigszins vóór de datum waarop het INSP expireerde lag (het eindrapport werd op 1 juli 1988 opgeleverd, terwijl het INSP formeel werd beëindigd op 31 december van dat jaar).

De fasering van activiteiten bij deze beperkte evaluatie van het INSP was als volgt:

- oriëntatiefase, waarbij stake-holders geïnterviewd werden en met de opdrachtgevers (te weten de ministeries van Algemene

Zaken, Economische Zaken, Onderwijs en Wetenschappen en Landbouw en Visserij) werd onderhandeld over de preciese opdrachtomschrijving (november '87 tot en met januari '88);
- formele instelling van de commissie voor de evaluatie van het INSP, en inschakeling van het bureau Horringa & De Koning als uitvoerend 'secretariaat' van de commissie;
- uitbesteding van de 6 deelstudies (maart '88);
- uitvoering van de deelstudies (april, mei '88);
- opstelling van de voorlopige versie van het eindrapport van de commissie, en bespreking van het concept met nationale en internationale panels van deskundigen (juni '88);
- oplevering eindrapport inclusief de rapporten van de deelstudies (30 juni '88).

Tot slot moet worden gewezen op de bijzondere organisatorische inkadering van deze evaluatie. De commissie kreeg nl. duidelijk de mogelijkheid om in onafhankelijkheid te opereren, hetgeen onder meer tot uitdrukking kwam in de lump-sum financiering, waarbij de commissie dus zelf kon beslissen over de allocatie van financiële middelen voor deelactiviteiten en de vrijheid had om zelf een ondersteunend apparaat te kiezen. Deze grote onafhankelijkheid – waaraan geen afbreuk werd gedaan door de toevoeging van een drietal ambtelijke waarnemers aan de commissie – verschilt aanmerkelijk van de organisatorische inkadering van grootschalige evaluatieprojecten in het onderwijs, waar het ministerie budget-controle heeft over ieder deelproject en er veelvuldig gewerkt wordt met ambtelijke secretarissen en ambtelijke stuurgroepen.

4.2 *Resultaten*

De voornaamste conclusies van de evaluatie van het INSP als geheel waren als volgt:
- De IT-ontwikkeling en de ingrijpende structurele maatschappelijke veranderingen die deze met zich meebrengt, bevinden zich nog in de eerste fase van de ontwikkeling.
- In onderdelen van het INSP komt de noodzakelijke accentverschuiving van een aanbod-gericht 'technology-push' beleid naar een beleid gericht op stimulering van grote groepen (potentiële) gebruikers niet of onvoldoende van de grond. Als deze vaak ook onvoldoende onderkende accentverschuiving niet alsnog wordt gerealiseerd, zullen de effecten van IT-stimulering lager uitvallen dan mogelijk was geweest.
- Door een gebrek aan samenhang tussen en binnen de verschillende onderdelen van het INSP zijn irritaties bij de doelgroepen ontstaan en zijn de effecten lager dan mogelijk zou zijn geweest.
- De twee hiervoor gemaakte opmerkingen hebben geen betrekking op de onderdelen binnen het landbouwgedeelte van het INSP. Deze onderscheiden zich in het algemeen door hun geïntegreerde, goed doordachte aanpak.
- Een groot aantal activiteiten op IT-gebied, in de sfeer van onderwijs, onderzoek, de landbouw en andere sectoren van het bedrijfsleven, was niet of trager van de grond gekomen zonder het INSP.
- De uiteindelijke effecten van belangrijke delen van het INSP zullen eerst in de loop van de komende jaren optreden. De omvang van die effecten zal sterk afhangen van de mate waarin de overheid en andere betrokken partijen op basis van de ervaringen tot nu toe hun activiteiten bijstellen. In die zin dienen belangrijke onderdelen van het INSP beschouwd te worden als investeringen, die bij goed management nog veel kunnen bijdragen aan de resultaten van de 'B.V. Nederland'.

De voornaamste aanbevelingen van de commissie waren:
1. Ga ook na het aflopen van het INSP, eind 1988, door met de stimulering van IT. Het grote strategische belang van IT voor de Nederlandse economie en de met IT-stimulering bereikte en nog te bereiken effecten rechtvaardigen een actief beleid op dit terrein.
2. Met erkenning van de bestuurlijke complexiteit van de overheidsorganisatie wordt aanbevolen in de post-INSP periode de strategievorming t.a.v. het IT-beleid te centraliseren en een en ander bestuurlijk te plaatsen in het kader van de voorstellen van de Commissie Dekker. Het is van cruciaal belang, dat bij deze strategieformulering aspecten van technologie, wetenschap en onderwijs geïntegreerd in

beschouwing worden genomen.
Meer in het bijzonder dienen de volgende functies op centraal niveau verricht te worden:

- keuze van de centrale thema's waarop de aandacht geconcentreerd gaat worden (bijv. bedrijfsoverschrijdende logistiek) en de prioriteiten tussen deze thema's;
- opsporen en signaleren van 'witte vlekken' op de terreinen van IT-onderzoek en toepassing;
- bewaking van de samenhang binnen de stimuleringsprogramma's;
- definiëring van de missies van de bij de uitvoering van het IT-beleid betrokken instituties;
- globale voortgangsbewaking van de deelplannen.

3. Decentraliseer de uitvoering van het IT-beleid naar instellingen met duidelijke missies per thema. Binnen deze missie moeten deze instellingen flexibel kunnen reageren op de snelle technologische en sociaaleconomische veranderingen door aanpassing van doelgroepen en policymix. Regelmatig wordt gerapporteerd over de bereikte resultaten tegen de achtergrond van concrete middellange termijnplannen met operationele doelstellingen en mijlpalen.
4. Leg binnen de geselecteerde thema's sterker het accent op stimulering vanuit de vraagzijde (kennisoverdracht, demonstratieprojecten, integratie van technologische, bedrijfseconomische, organisatorische, management- en trainings-aspecten). Geef daarbinnen prioriteit aan het oplossen van opleidingsvraagstukken.
5. Plaats het IT-beleid meer in internationale kaders.
6. Geef extra aandacht aan wet- en regelgeving met betrekking tot standaardisatie/normalisatie, privacy, auteursrechten, aansprakelijkheid, enz.
7. Zorg voor continuïteit op deelterreinen, zoals de stimulering van het irregulier onderwijs, nadat het INSP formeel is afgelopen en de hierboven voorgestelde structuur nog niet is geïmplementeerd.

Wat het onderwijsdeel van het INSP betrof, formuleerde de evaluatiecommissie vier hoofdconclusies:

- Bij alle toepassingsvormen van de computer in het onderwijs is gebleken, dat de daadwerkelijke invoering een geïntegreerde aanpak vraagt van hard- en softwarevoorziening, inbedding in het curriculum en (na-)scholing van docenten. Een dergelijk veranderingsproces kost veel tijd en is moeilijk te bestuderen. De resultaten tot nu toe lijken in overeenstemming te zijn met deze moeilijkheidsgraad.
- Het NIVO-project, uitgevoerd in samenwerking met het bedrijfsleven, levert een belangrijke bijdrage aan de bekendheid met de computer in het voortgezet onderwijs in het kader van het vak informatiekunde en aan de standaardisatie.
- De aansluitingsproblematiek tussen onderwijs en arbeidsmarkt blijft een majeur probleem. Reguliere onderwijsinstituten missen de informatie en de flexibiliteit om snel en gericht op veranderingen in functies en beroepen te kunnen inspelen. Het irreguliere onderwijs speelt evenmin systematisch op de aansluitingsproblematiek in, maar beschikt wel over de flexibiliteit die het reguliere systeem ontbeert. De regionale instituten dreigen door een onduidelijke missie tussen de wal (beroepsonderwijs) en het schip (irregulier onderwijs) te vallen.
- Met het POCO-project beschikt Nederland voorlopig over een bevredigende organisatievorm voor courseware-ontwikkeling. De daadwerkelijke toepassing van COO binnen de diverse onderwijscategorieën heeft nog een incidenteel karakter.

De aanbevelingen die de commissie deed ten aanzien van toekomstige IT-stimulering in het onderwijs luidden als volgt:

- Geef veel aandacht aan de hierboven in algemene zin bepleite afstemming met het technologie- en wetenschapsbeleid.
- Beperk de missie van het reguliere beroepsonderwijs tot de basisvaardigheden, stimuleer irreguliere vormen van onderwijs en koppel deze twee systemen via de regionale instituten, vormen van contractonderwijs e.d.
- Zorg ervoor dat het project-management binnen de OPSTAP-structuur met voldoende daadkracht en autonomie kan opereren.
- Geef hogere prioriteit aan onderwijsresearch op het gebied van IT.

De resultaten van de twee deelstudies die betrekking hadden op het deelgebied onderwijs van het INSP, nl. de studies betreffende courseware en IT-stimulering in regulier en irregulier beroepsonderwijs, zullen hier niet worden samengevat. Voor de resultaten van de laatstgenoemde deelstudie zij verwezen naar Nijhofs bijdrage aan dit themanummer; voor wat betreft de eerstgenoemde deelstudie kan worden verwezen naar de rapportage van de evaluatiecommissie INSP (1988).

5 *Discussie*

Het bijzondere van de evaluatie van het onderwijsdeel van het INSP als onderwijsevaluatieproject was dat deze deel uitmaakte van een integrale evaluatie van het INSP. Een inhoudelijke uitkomst die zonder deze onderbrenging in een ruimere evaluatie van informaticastimulering nooit zou zijn vastgesteld, was de constatering dat in alle niet speciaal op het onderwijsdeel betrekking hebbende deelstudies (produktie-automatisering, software-sector, bedrijfsoverschrijdende logistiek en complexe systemen) opleiding en scholing als een hoogst belangrijk aspect werd onderkend. De evaluatie-uitkomsten over de niet-onderwijs onderdelen van het INSP bevestigen dus het belang van informaticastimulering in het onderwijs.

Een tweede belangrijke uitkomst was de constatering dat er voor het onderwijsdeel van het INSP weliswaar een meer geprononceerde managementstructuur bestond dan voor de meeste andere onderdelen, maar dat ook voor dit onderdeel het algemene patroon van een sterk gefragmentariseerd geheel van deelprojecten en -projectjes opging. Kennelijk was dus ook een managementstructuur, die naar Nederlandse verhoudingen op het terrein van onderwijsvernieuwing een betrekkelijk 'top-down'-karakter had, geen waarborg voor een sterk geïntegreerde beleidsuitvoering. Het is echter een open vraag of dit beter zal gelukken bij de projectorganisatie van onderwijsverzorgingsinstellingen die thans verantwoordelijk is voor het vervolg van het onderwijsdeel van het INSP. De evaluatiecommissie heeft op dit punt – overigens ten aanzien van het gehele INSP – de aanbeveling gedaan van een sterk centralistische aanpak op het niveau van strategiebepaling en uitvoeringsinspectie en een gedecentraliseerde operationele planning door daartoe in aanmerking komende professionele organisaties. Het zou hier te ver voeren om nader in te gaan op de bijzondere kenmerken van het onderwijsdeel van het INSP als onderwijsinnovatieproject, hoewel het project als zodanig bijzonder interessant is. Ook vanuit het gezichtspunt van de organisatie van grootschalige onderwijsevaluatieprojecten was de hier centraal staande evaluatie bijzonder. In het voorafgaande is al gewezen op de grotere mate van onafhankelijkheid ten opzichte van de betrokken departementen dan doorgaans het geval is bij grootschalige onderwijsevaluaties. Een tweede opmerkelijk kenmerk van de organisatie van dit evaluatieproject was de ruime inschakeling van particuliere organisatie-adviesbureaus, zowel bij de directe ondersteuning van de commissie als bij de uitvoering van de deelstudies. Beide kenmerken (onafhankelijkheid van de overheidsbureaucratie en inschakeling van particuliere bureaus) hebben in belangrijke mate bijgedragen aan de realisering van dit omvangrijke evaluatieproject (met een totaal budget van 1 miljoen gulden) in een periode van krap een half jaar – alle rapporten van de deelstudies en het eindrapport van de commissie werden op de vooraf overeengekomen data opgeleverd. De onafhankelijkheid van de uitvoeringsorganisatie van dit evaluatieproject betekende overigens niet dat er niet intensief met de opdrachtgevers, zijnde de betrokken departementen, werd samengewerkt. Zo waren er onder meer veelvuldige contacten om te komen tot een taakomschrijving van de commissie, werd er constructief samengewerkt met de ambtelijke waarnemers in de commissie en werden de uitkomsten van de evaluatie doorgenomen met topambtenaren van de betrokken departementen.

Ook vanuit methodologisch gezichtspunt was de INSP-evaluatie ongebruikelijk. Dit gold om te beginnen voor de toepassing van een evalueerbaarheidsanalyse die vooraf ging aan (de eerste fase van) de eigenlijke evaluatie. De 'sobere, toekomstgerichte' evaluatie was een mengeling van analytische en empirische benaderingen, waarbij een zware wissel getrokken werd op expertbeoordelingen. De manier waarop de evaluatie werd uitgevoerd had zekere overeenkomsten met evaluatiebenaderingen die in de literatuur bekend staan als 'stake-holder based evaluation' (Cohen,

1983) en met 'utilization focused evaluation' (Alkin, Daillak & White, 1979; Patton, 1978). De overeenkomst met deze benaderingen kwam bijvoorbeeld tot uitdrukking in het consulteren van vertegenwoordigers van allerlei maatschappelijke groeperingen bij de keuze van thema's en invalshoeken bij de evaluatie, de tijdige oplevering van rapporten, en het terugkoppelen van gegevens naar sleutelfiguren op tijdstippen waarop deze gegevens gebruikt konden worden voor besluitvorming. De intentie van de commissie, om vooral te letten op hoofdlijnen heeft met zich meegebracht dat de evaluatie-uitkomsten vaak een nogal algemeen karakter hebben. Kritische kanttekeningen die bij een dergelijke evaluatiebenadering te plaatsen zijn, hebben dan ook vooral betrekking op de 'hardheid' en specificiteit van de onderzoeksuitkomsten. Extreem gesteld heeft deze evaluatie meer weg van een onderbouwd advies dan van een traditionele empirische evaluatie. Ter verdediging op deze kritiek kan worden aangevoerd dat gezien de complexiteit van het evaluatieobject en de korte tijd die voor het onderzoek beschikbaar was een meer gedetailleerde en sterker empirische evaluatie eenvoudigweg niet uitvoerbaar was.

Noten

1. De commissie voor de analyse van de evalueerbaarheid van het INSP bestond uit prof. dr. J. S. Cramer, prof. dr. J. Scheerens en prof. ing. W. C. L. Zegveld. De commissie die de beperkte formatieve evaluatie uitvoerde bestond uit prof. ing. W. C. L. Zegveld (voorzitter), prof. dr. J. Scheerens en mr. T. L. Steehouwer. Deze commissie werd ondersteund door het bureau Horringa & De Koning.

Literatuur

Alkin, M. C., R. Daillak & P. White, *Using evaluations*. Beverly Hills: SAg Library of Social Research, 1979.

Beem, A L., (red.), *Weten, redeneren, raden. Een analytische evaluatie van de Basisvorming*. Lisse: Swets & Zeitlinger, 1988.

Beem, A L., Y. J. Pijl & J. Scheerens, *Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma Onderwijsvoorrangsbeleid*. Den Haag: SVO, 1986.

Cohen, D. K., Evaluation and reform. In: *Stakeholderbased evaluation* (New Directions for Program Evaluation no. 17) (pp. 73-81). San Francisco: Jossey Bass, 1983.

Commissie Evaluatie Informaticastimuleringsplan, *Een eerste aanzet. Verslag van een sobere toekomstgerichte evaluatie van het Informaticastimuleringsplan*. Den Haag: Staatsdrukkerij, 1988.

Hofstee, W. K. B., J. Schaveling & K. B. Koster, *Beleidsanalytische verkenning. Onderwijsvoorrangsbeleid op schoolniveau*. Groningen: Rijksuniversiteit Groningen, 1987.

Hoogerwerf, A., Beleid berust op veronderstellingen: de beleidstheorie. *Acta Politica*, 1984, *19*, 493-531.

Jung, S. M. & J. G. Schubert, Evaluability assessment: A two year retrospective. *Educational Evaluation and Policy Analysis*, 1983, *5*, 435-444.

Patton, M Q., *Utilization-focused evaluation*. Beverly Hills: Sage, 1978.

Plomp, Tj., K. van Deursen & J. Moonen (Eds.), *CAL for Europe*. Amsterdam: North-Holland Publishing Comp., 1987.

Scheerens, J., Contextual influences on evaluations: the case of innovatory programs in Dutch education. *Educational Evaluation and Policy Analysis*, 1985, *7*, 309-317.

Curriculum vitae

J. Scheerens is als hoogleraar verbonden aan de Universiteit Twente, Faculteit Toegepaste Onderwijskunde. Zijn leeropdracht ligt op het terrein van onderwijsorganisatie en -management. Hij maakte deel uit van verschillende evaluatiecommissies en is momenteel voorzitter van het Interuniversitair Centrum voor Onderwijsevaluatie (ICO): een samenwerkingsverband voor 2de fase onderwijs van de Universiteit Twente en de Rijksuniversiteit Groningen.

Adres: Universiteit Twente, Postbus 217, 7500 AE Enschede

Manuscript aanvaard 6-12-'89

Summary

Scheerens, J. 'The evaluation of the educational section of the Informatics Stimulation Program.' *Pedagogische Studiën*, 1990, *67*, 88-99

In this article the context, methodology and results of two evaluation studies concerning the Dutch Informatics Stimulation Programme (INSP) are discussed, the educational section of the programme being the main focus of interest. Of the two evaluation studies that are discussed, the first was an evaluability analysis, and the second was a "sober, future oriented evaluation". Each study was conducted within a period of roughly six months. The strengths and weakenesses of the particular evaluation approach that was used in the second study are discussed, the conclusion being that the proactive, improvement oriented and partly analytic (as well as empirical) nature of this study was justified by the stage of development of the programme to be evaluated and by contextual constraints. Substantive findings that were most striking are the recognition of importance of education and training for all elements of IT-stimulation in society and the seriousness of the problem of how to organize and control large-scale IT-stimulation in education.

Vliegwieleffecten van een stimuleringsprogramma
Een evaluatie van INSP-deelprogramma's gericht op het beroepsgeoriënteerde onderwijs

W. J. NIJHOF
Vakgroep Curriculumtechnologie, Universiteit Twente

Samenvatting

Het Informatica Stimuleringsplan (INSP) dat effectief was in de jaren 1984 t/m 1988 heeft vooral gefunctioneerd als een prikkel om initiatieven te bevorderen. Binnen dit plan was een deelprogramma gericht op het beroepsgeoriënteerde onderwijs. Dat deel is voorwerp geweest van een bliksemevaluatie. De conclusie van die evaluatie is dat het programma heeft gewerkt als een vliegwiel en een diversificatie aan verwachte en onverwachte effecten heeft gesorteerd die niet allemaal even gunstig uitpakken. De belangrijkste conclusie is dat in de aanvangsfase heel naief de nadruk lag op apparatuurverstrekking, dat er geen samenhang was met leerplanontwikkeling en nascholing en dat de beoogde kwalificatieontwikkeling, nl. aansluiten op verandering in de beroepenstructuur niet uit de verf is gekomen. Verwacht wordt dat een aantal vliegwieleffecten in het vervolgprogramma PRINT beter tot zijn recht komt. Een echte output-evaluatie is dan wèl op zijn plaats.

1 *Inleiding en probleemstelling*

Eén van de hoofddoelstellingen van het INformatica StimuleringsPlan (INSP) (1984-1988) was het creëren van 'human capital' op de terreinen informatiekunde en informatica. Dit doel werd binnen het overheidsbeleid als noodzakelijk gezien, omdat bij het bedrijfsleven en ook binnen de politiek de zorg bestaat dat het reguliere beroepsonderwijs onvoldoende in staat is om op deze terreinen tegemoet te komen aan de behoeften op de arbeidsmarkt.

Een van de vragen die het beleid zich o.a. stelt is in hoeverre het niet-reguliere onderwijs in staat moet worden geacht deze aansluiting wèl tot stand te brengen. Onder niet-regulier onderwijs wordt in dit verband verstaan particuliere opleidingen die worden aangeboden op de markt en waaraan in principe iedere volwassene kan deelnemen. Er zijn ook (interne bedrijfs-)opleidingen die worden aangeboden door werkgevers en uitsluitend toegankelijk zijn voor het personeel van de betreffende onderneming of instelling (zie Mulder, Akkerman & Bentvelsen, 1988).

Bezien dient echter te worden of de stelling juist is dat het reguliere beroepsonderwijs niet in staat zou zijn adequaat en flexibel in te spelen op de behoeften van de arbeidsmarkt. Bovengenoemde vraagstelling gaat te gemakkelijk uit van een tweedeling en van de veronderstelling dat wat het reguliere systeem (nog) niet kan of mag wel geschiedt door het private onderwijssysteem.

Verschillende onderzoeksprojecten zijn de laatste jaren ontwikkeld en uitgevoerd om te onderzoeken hoe de aansluiting onderwijs-arbeidsmarkt m.b.t. informatietechnologische ontwikkelingen gerealiseerd kan worden (o.a. Van Hoof, 1986; Hesse & Nijhof, 1988; Kanselaar, De Bruijn, Cadot, Dehning, Groot & Maas, 1987; Nijhof & Mulder, 1986; Pieters & Mulder, 1988). Daarbij moet vastgesteld worden dat het om een relatief nieuw domein gaat waarvoor noch in het buitenland, noch in ons land langdurige routines en ervaringen zijn opgebouwd. Bovendien is de basis voor de ontwikkelingen van nieuwe opleidingen niet altijd aanwezig. Deze bestaat doorgaans uit een gekend en bekend beroepsprofiel. Door de nieuwe technologische ontwikkelingen ontstaan daarin (soms ingrijpende) veranderingen. Een heroriëntatie treedt daardoor op, zowel binnen bedrijven als binnen opleidingen. Van daaruit is het begrijpelijk dat de vraag naar de functie en de functionaliteit van het reguliere en het niet-reguliere beroepsgerichte systeem aan de orde is waar het de aansluitingsproblematiek

Pedagogische Studiën 1990 (67) 100-115

betreft. Deze aansluitingsproblematiek wordt in de context van dit artikel vooral vanuit een leerplantechnische optiek bezien: in welke mate slagen opleidingen erin nieuwe kwalificatiestructuren aan te reiken resp. te ontwikkelen op het domein van de informatiekunde binnen de beroepsgerichte opleidingen in het licht van de veranderende arbeidsmarkt. De termen informatica en informatiekunde vatten wij in dit verband op als kennisgebieden of leerplandomeinen, terwijl onder informatietechnologie de kennis wordt verstaan die gebruikt wordt om informatie te be- en verwerken d.m.v. informatieverwerkende systemen.

De commissie Evaluatie INSP (zie de bijdrage van Scheerens in dit themanummer) heeft aan bovengenoemd aansluitingsvraagstuk een aantal aspecten onderscheiden, en wel:

a. de discrepantieproblematiek tussen onderwijsaanbod en de behoeften op de arbeidsmarkt;
b. de mogelijkheden van het reguliere en niet-reguliere beroepsgerichte onderwijs, incl. de innovatiecapaciteit van beide in vergelijkend opzicht;
c. stimulering van het irreguliere 'stelsel' voor de bevordering van informatiekunde resp. informatica.

Aan de hand van een uiterst sobere achteraf-en een op de toekomst gerichte evaluatie diende een antwoord te worden verkregen op de bovengestelde vraagstellingen. Het sobere karakter van de evaluatie bestond daarin dat slechts 20 werkdagen beschikbaar waren en een totaal bedrag van *f* 40.000,– . De evaluatie moest daardoor het karakter krijgen van een bliksemonderzoek en voortbouwen op beschikbare kennis.

Een van de onderdelen van de evaluatie INSP heeft betrekking op het lager en middelbaar beroepsonderwijs (INSP cluster II.4). De doelstelling van dat cluster luidt: "Bevorderen c.q. verder uitbouwen van aspecten van informatietechnologie in het leerplan van de verschillende schoolsoorten" (Informaticastimulering, oktober 1987, p. 10). Na aanvankelijk binnen het INSP het accent te hebben gelegd op de apparatuurverstrekking, lijkt de aandacht verschoven te worden naar leerplanontwikkeling en programmatuurontwikkeling voor de sectoren techniek, economie/administratie en dienstverlening (gezondheidszorg). De evaluatie is daarom vooral gericht op de vraag:

a. of 'human capital' wordt gecreëerd (kwalificaties);
b. of discrepanties tussen onderwijsaanbod en behoeften op de arbeidsmarkt zijn afgenomen;
c. wat de innovatiecapaciteit van regulier en niet-regulier beroepsgericht onderwijs is;
d. welke vormen van stimulering noodzakelijk zijn voor de bevordering van informatiekunde resp. informatica.

De antwoorden hebben betekenis voor het stimuleringsbeleid inzake toepassingen van informatietechnologie in het onderwijs na het Informatica stimuleringsplan o.a. in het Programma Invoering Nieuwe Informatie Technologieën in het Onderwijs (PRINT).

2 *Context*

2.1 *Het Cluster II.4 (LBO/MBO-projecten)*

Voor een goed begrip van cluster II.4 wordt hier zeer beknopt een schets gegeven van de vragen, overwegingen en criteria die hebben geleid tot keuzen voor het inhoudelijk beleid. Bij de formulering van het inhoudelijk kader voor het projectplan speelde een viertal kernvragen een rol, te weten:

1. Waar zijn in het bedrijfsleven, in de dienstverlenende sector en bij de overheid belangrijke ontwikkelingen te verwachten?
2. In welke mate speelt de introductie van informatica in welke toepassingsvorm dan ook daarbij een rol?
3. Vanuit welke schoolsoort vindt een belangrijke uitstroming plaats naar deze delen van het bedrijfsleven en de dienstverlening?
4. Met welke activiteiten moet door de scholen op deze ontwikkelingen worden ingespeeld (PSOI, 1987a, p. 3)?

Vooral vraag 4 wordt van grote betekenis geacht (PSOI, 1987a, p. 5).

Waar het de eerste drie vragen betreft is aangesloten bij conclusies van de 'Commissie Wagner II'. Deze leiden vervolgens tot zogenaamde hoofdaandachtsgebieden en 8 manifeste toepassingsgebieden van informatica, te weten:

1. Transito-informatica toepassing (transport, opslag, overslag en transito-activiteiten).

2. Industriële automatisering (CAD/CAM, Flexibele Produktie-Automatisering (FPA)).
3. Communicatieve informatieverzorgende systemen (o.a. databanken).
4. Systemen voor geavanceerde toepassing in bank- en verzekeringswezen.
5. Systemen voor electronic publishing.
6. Verkeersbegeleidingssystemen (wegverkeer, scheepvaart en luchtvaart).
7. Medische informatica en automatiseringssystemen.
8. Kantoorsystemen.

Op grond van bovenstaande vragen en overwegingen is een beleid geformuleerd dat op de volgende uitgangspunten is gebaseerd

a. het vernieuwingsbeleid binnen scholen en binnen cluster II.4 dient op elkaar te worden afgestemd;
b. beroepsvoorbereidend onderwijs moet geënt zijn op feitelijke in de beroepspraktijk voorkomende situaties;
c. bijstellen van leerplannen dient in kortlopende projecten te worden gerealiseerd;
d. streven naar sectorale samenhang tussen projecten;
e. experimenteel beproeven van optimaal gebruik van duurdere apparatuur b.v. door regionale voorzieningen;
f. beproeven van de relatie tussen nieuwe technologieën in bedrijfsleven en de leerplannen in het beroepsonderwijs. De aansluiting MBO-HBO dient zorgvuldig bestudeerd te worden.

Centraal staat de afstemmingsvraag en de samenhang tussen onderwijssoorten. De ontwikkeling van kwalificaties dient gebaseerd te zijn op feitelijk in de beroepspraktijk voorkomende zo mogelijk verwante beroepssituaties. Hoewel dit uitgangspunt aantrekkelijk lijkt blijkt uit nogal wat studies dat de ontwikkeling van de beroepenstructuur onder invloed van moderne technologieën allerminst eenduidig is en sterke polarisatietendenzen vertoont (Rumberger, 1987; Wegmann, Chapmann & Johnson, 1989; Spenner, 1985).

In de tweede plaats wordt een samenwerking tussen bedrijfsleven, regulier en niet-regulier onderwijs voorgestaan o.a. door middel van regionaal in te richten centra.

In de derde plaats zoekt men naar korte termijn aanpassingen van leerplannen om doeltreffend te reageren op ontwikkelingen in de arbeidsmarkt. Lange termijn trajecten zijn uit den boze. Gezocht wordt naar een verkorting van procedures t.b.v. leerplanontwikkeling en -invoering.

Gekozen is voor zowel een diepte- als een breedtestrategie. De dieptestrategie is gericht op het laten beproeven resp. ontwikkelen van programmatuur en courseware in nauwe relatie tot de aangeschafte hardware en wordt betaald uit INSP-gelden. De breedtestrategie is gericht op de verspreiding van gevonden oplossingen binnen cluster II.4 en wordt gefinancierd uit reguliere middelen. Het totale financiële INSP-kader voor cluster II.4 bedroeg ca. 75 miljoen gulden (1984-1988).

Om projecten en prioriteiten te kunnen bepalen die binnen het programma passen zijn de volgende overwegingen gehanteerd:

a. Er is behoefte aan *concrete voorbeelden van toepassingen van informatietechnologie* (IT) in de beroepspraktijk. Op basis daarvan kunnen heldere en voor langere perioden geldige leerplannen worden ontwikkeld (PSOI, 1987a, p. 9).
b. Voor de invoering van IT zijn *educatieve produkten* noodzakelijk. Primair zal gezocht worden naar in het buitenland aanwezige produkten.
c. Qua *software* dient gebruik te worden gemaakt van produkten die in de beroepspraktijk worden benut. Uniformering wordt voorgesteld. Problemen m.b.t. auteursrechten en aanpassingen dienen te worden onderzocht.
d. *Hardwareconfiguraties* dienen te voldoen aan: 16 bits architectuur, werkend onder MSDOS. De verstrekkingsbasis is 8 computers per school.
e. Het vak *informatiekunde* zal als gevolg van ontwikkelingen in het v.o. (NIVO-project) moeten worden omgebouwd in *een beroepsgerichte richting*. Dit geldt voor lbo, kmbo, bbo/leerlingwezen en mto (Computerkunde).

De projecten binnen Cluster II.4 zouden ten slotte moeten voldoen aan de volgende criteria (PSOI, 1987a, 14): Ze moeten passen in het geformuleerde beleid, innovatief zijn, planmatig zijn opgezet, concreet gedefinieerde overdraagbare produkten opleveren, een hanteerbare omvang en een doorlooptijd van maximaal 3 jaar hebben, een goede projectorganisatie en een kwalitatief goede bemanning

garanderen en de mogelijkheid bieden voor inbreng van extern aanwezige expertise. Ondoelmatige overlap van projecten dient te worden tegengegaan. De bewaking van de voortgang berust bij het project-management INSP-cluster II.4.

Ter complicering van het geheel werd de commissie Evaluatie INSP door de overheid i.c. de Ministeries van Landbouw en Visserij en Economische Zaken gewezen op twee deelprojecten die weliswaar geen deel uitmaakten van cluster II.4, maar qua thematiek wel verwant waren en ook onderdeel waren van het INSP, te weten de Informatisering in de landbouw en het landbouwonderwijs en het Nationaal Inhaalprogramma Informatica Opleidingen (NIIO). Omdat ook deze beide projecten voorwerp werden van evaluatie worden ze hier beknopt beschreven.

2.2 *Informatisering in de landbouw en het landbouwonderwijs*

Om de automatisering in de landbouw te bevorderen heeft het Ministerie van Landbouw en Visserij een eigen INSP-LV opgezet. Doelstelling van dat plan was het behouden en versterken van de internationale concurrentiepositie van de agrarische bedrijfstak o.a. door een goed gebruik te maken van informaticatoepassingen. Het zwaartepunt van het INSP-LV ligt bij een gerichte stimulering van de automatisering van boerderij- en managementsystemen. Binnen het INSP-LV is gewerkt aan:

a. de opbouw van zogenaamde takorganisaties en het uitvoeren van proefprojecten;
b. een sterke combinatie van onderzoek, onderwijs en voorlichting;
c. een informatica-infrastructuur van het ministerie van L & V.

Van groot belang in het kader van dit INSP-LV plan was het ontwikkelen van informatiemodellen.

Voor het plan was in totaal 236 miljoen gulden beschikbaar, waarvan een deel benut is voor landbouwkundig onderzoek, waartoe proefboerderijen en regionale onderzoekscentra zouden moeten fungeren in het gebruik van managementsystemen.

T.b.v. het informatica-onderwijs op hogere, middelbare en lagere agrarische scholen (HAO, MAO en LAO) is veel aandacht besteed aan de coördinatie tussen schoolsoorten en de afstemming met bedrijfsleven. Gekozen is voor de volgende fasering:

Fase 1: eerst maken van leerplannen voor basisinformatica; in proefprojecten wordt de ontwikkelde programmatuur beproefd en wordt aan docenten nascholing gegeven.

Fase 2: gebruik van informatica bij de lessen in de beroepsgerichte en ondersteunende vakken. Dit gebeurt in overleg met de takorganisaties. Ook wordt bijscholing verzorgd voor werkenden in de agrarische sector. Voorts zijn samenwerkingsprojecten opgezet tussen agrarische hogescholen in het kader van levensmiddelentechnologie, bosbouw en cultuurtechniek (IBM, 1987).

Alle informatica-ondersteunings- en coördinatie-activiteiten zijn ondergebracht bij het Centrum voor Informatica Ondersteuning voor het Landbouwonderwijs (CILO). Binnen deze organisatie wordt gewerkt aan de ontwikkeling van programmatuur, handleidingen en aan bijscholing. Het CILO heeft daarmee een centrale taak binnen het agrarisch onderwijs. Het verzorgt niet alleen de courseware, maar ook de implementatie, voorlichting, opleiding individuele begeleiding, onderhoud en applicatie- en systeembeheer.

Qua apparatuur beschikt het agrarisch onderwijs vrijwel volledig over bestuurlijk-administratieve voorzieningen. Voor onderwijskundig gebruik zijn nog geen grote ontwikkelingen te melden.

Het beschikbaar stellen van apparatuur is een van de knelpunten binnen het INSP-LV programma. Waarmee de toepassing van informaticakennis uiteraard wel staat of valt. Daarmee wordt tevens vastgesteld dat dit project geen technology push als startfase kende.

2.3 *Het Nationaal Inhaalprogramma Informatica Opleidingen (NIIO)*

Binnen het kader van het INSP zijn projecten georganiseerd die betrekking hebben op zowel regulier als niet-regulier onderwijs. Het Ministerie van Economische Zaken (EZ) was met name betrokken bij projecten gericht op het beroepsonderwijs. INSP-projecten voor het regulier onderwijs zijn verdeeld volgens de onderwijssectoren, terwijl de irreguliere projecten in vijf clusters zijn ondergebracht.

Het NIIO dat onder de hoede van EZ verkeerde heeft als doel de kwalitatieve en kwantitatieve discrepantie tussen vraag en aanbod voor een aantal beroepsgroepen, die van

belang zijn voor de stimulering van het informaticagebruik, te verminderen. Het NIIO bestond uit de volgende vijf clusters:

1. Bijscholing van werkzame academici.
2. Bijscholing van overheidsmanagers.
3. Praktijktrainingscentra voor produktievernieuwing (PTP); Automatiseringspraktijkcentra (APC's).
4. Demonstratieproject voor het midden en klein bedrijf (mkb).
5. Informatica in de transport- en transitosector (INTIS).

Voor het programma-onderdeel was 50 miljoen gulden beschikbaar.

3 *Evaluatieopzet*

T.b.v. de evaluatie van het INSP was voorzien in een evaluatieplan (Scheerens, 1989). Doel van dat plan was een oordeel te vellen over de generale effecten van het INSP dat qua projectorganisatie en financiële omvang uniek was. Hoewel een evaluatie voorzien was, afgedwongen door de Tweede Kamer, kon deze slechts aan het einde van het vijfjarig programma worden uitgevoerd. Evaluaties die eigenlijk vooraf gepland behoorden te worden, zoals tussentijds door externe evaluatoren, om het zicht op de voortgang en de effecten te behouden waren niet aanwezig.

Gedwongen door tijd en de beperkte hoeveelheid geld kreeg de evaluatie van het INSP daardoor het karakter van een noodzakelijke, maar beperkte sondering, waarvan de resultaten mogelijkerwijs niet meer zouden kunnen worden benut in voortgangsplannen, zoals het PRoject Invoering Nieuwe Technologieën (PRINT). Deze overheidsplannen voor de verdere integratie van informatietechnologie in het onderwijs waren medio 1988 reeds nagenoeg vastgesteld.

De vraag was waartoe deze evaluatie zou moeten dienen. Een niet zo bijzondere vraag t.a.v. overheidsevaluaties. De maximale verwachting zou in ons geval kunnen zijn een beter inzicht in de opbrengsten en wellicht een marginale bijsturing van overheidsplannen. Evaluatie-onderzoek als stiefkind derhalve.

Toch is er nog een specifieke reden. Het beroepsonderwijs is een sector die in het bijzonder geconfronteerd wordt met de gevolgen van informatietechnologische toepassingen in beroep en bedrijf die doorgaans direct worden vertaald in vaardigheden en bekwaamheden die noodzakelijk zijn om het be- en verwerkingsproces van produkten van welke aard en soort adequaat te laten verlopen. Anders dan in het algemeen voortgezet onderwijs is de opleiding toegesneden op beroepsgerichte vaardigheden, al is er een sterke behoefte om meer algemene op transfer gerichte 'basisvaardigheden' te onderscheiden en deze op te nemen in leerplannen (zie b.v. Nijhof & Mulder, 1986; Nijhof & Remmers, 1989). Cluster II.4 is met name gericht op de vraag of de aansluiting met de beroepssituatie in acht wordt genomen. Evaluatie van dit fenomeen alleen al levert een belangrijk inzicht in de wijze waarop scholen en projecten dit probleem te lijf gaan.

Zinvol is vooral de vraag of via de dieptestrategie structurele oplossingen worden gevonden en of deze dissemeneerbaar zijn.

Het voorgaande maakt duidelijk dat de voorgestelde evaluatie niet het karakter had van een formatieve of summatieve leerplanevaluatie, maar van een zeer sobere beleidsevaluatie op enkele hoofdpunten waar het de aansluitingsproblematiek betreft: kwantitatief en inhoudelijk en waar het betreft de vormen van samenwerking en ontwikkeling tussen regulier en niet-regulier beroepsgericht onderwijs.

Vergelijking van 'beide' groepen is weliswaar een doel, maar niet eenvoudig vanwege de verscheidenheid aan doelstellingen en objecten. Voor het niet-reguliere onderwijs lag geen plan voor en waren ook geen bijzondere faciliteiten beschikbaar.

De selectie van gevallen

De selectie van projecten (n = 14) uit het totaalbestand (n = 60) is tot stand gekomen in overleg met de Commissie INSP-evaluatie en een groep beleidsambtenaren van O & W, L & V en EZ. De projecten moesten voldoen aan de volgende eisen:

- de clusterdoelstellingen fungeren als kader;
- de selectie vormt een dwarsdoorsnede van het beroepsonderwijs;
- een accent wordt gelegd op het technisch onderwijs;
- een aparte aanvulling vormt het Informatiseringsprogramma van het Landbouwonderwijs.

Daarnaast wordt aandacht besteed aan het

Nationaal Inhaal programma Informatica Opleidingen (NIIO) en aan andere niet-reguliere opleidingen. Dit betekent dat de lijst met projecten een select karakter heeft.

Toepassing van bovenstaande criteria heeft geresulteerd in het overzicht dat is weergegeven in Tabel 1.

Tabel 1 *Te evalueren INSP-projecten**

1. Speerpunt experiment regionale centra CNC/MTO
2. Micro electronica MTO
3. Besturingstechniek MTO
4. Computer Aided Design (CAD) MTO
5. Flexibele automatisering m.b.v. lasrobots
6. Fotozetsystemen grafische MTS'en
7. Experimentele technische computerkunde MTO
8. Proefproject Robotica in het KMBO
9. Uitbreiding technische computerkunde MTS
10. SOMprojecten CNC/CAD
11. Ontwikkelingen VEV/BKA
12. Speerpuntproject electrotechniek
13. Technische informatica Bovenbouw LTO
14. Voorbeeldcentra Kantoor- en Winkelpraktijk
15. Het Informatiseringsproject LV
16. Het Nationaal Inhaalprogramma Informatica Opleidingen
17. Het niet-reguliere onderwijs (particulier)

* Projectomschrijvingen zijn te vinden in PSOI-reeks, nr. 27, 1987b.

Methoden en technieken

Twee hoofdinstrumenten zijn benut om gegevens te verwerven: interviews en documentenanalyse.

Een tiental respondenten is benaderd voor overzichtsinformatie en het geven van een zicht op de toekomst. Het betreft hier beleidsambtenaren van de verschillende departementen, wetenschappelijke experts en INSP-projectmanagers. Waar het het Nationaal Inhaalprogramma Informatica Opleidingen betreft is volstaan met gesprekken met beleidsambtenaren van het departement van Economische Zaken en zijn enkele bronnen geraadpleegd. Voorts zijn van alle projecten de projectcoördinatoren benaderd voor een interview. In totaal ging het om ca. 35 personen. Door het feit dat de respondenten, projectcoördinatoren, projectmanagers of beleidsambtenaren, betrokken waren bij de ontwikkeling van de plannen en projecten is de kans op antwoordtendenties groot. Aan de hand van eigen waarneming, externe evaluatierapporten en analyse van documenten van de projecten is deze informatie gecontroleerd, gecorrigeerd of aangevuld.

Om op een systematische wijze gegevens te verzamelen zijn zogenaamde topiclijsten ontwikkeld (zie Tabel 2). Deze lijsten bieden de mogelijkheid alle gewenste informatiecategorieën na te lopen en toch ruimte te laten voor niet voorziene ontwikkelingen.

Tabel 2 *De topiclijst*

0. Inhoud van het project
1. Doelstelling
2. Apparatuur (omvang, standaardisatie, kwaliteit, beheer, service)
3. Software (omvang, aard, kwaliteit, adequaatheid, beheer)
4. Nascholing (omvang, door wie)
5. Ervaringen nascholing (goed, matig, slecht)
6. Leerervaringen Docenten
7. Leerervaringen Studenten
8. Leerplanontwikkeling
9. Integratie van IT in beroep
10. Integratie IT in leerplannen
11. Integratie IT in de onderwijsorganisatie
12. Aansluiting IT op voorgaand onderwijs
13. Inpassing in SVM-operatie
14. Afstemming met niet regulier onderwijs
15. Wijze van instructie
16. Combinatie van IT met andere technologieën
17. Combinatie van IT met moduleren

Alle gesprekken zijn op tape vastgelegd en schriftelijk verwerkt. De informatie uit documenten en gesprekken is verwerkt aan de hand van bovenstaande categorieën, uiteraard voorzover van toepassing. De evaluatie heeft daardoor het karakter gekregen van een vergelijkende case studie. De belangrijkste bevindingen per project zijn weergegeven in Nijhof, Reijmerink en Thijsen (1988, p. 32 e.v.).

In het kader van de evaluatie van het particulier onderwijs is een keuze gemaakt voor een vijftal grote tot zeer grote instituten die opleidingen verzorgen ten aanzien van toepassingen van informatietechnologie in beroep resp. functie, en wel: Dirksen, LOI, NOVI, PBNA en TELEAC. Deze instituten zijn benaderd voor schriftelijke informatie en er zijn gesprekken gevoerd met coördinatoren/respondenten m.b.t. de inspanningen die deze instellingen leveren t.a.v. scholing en

training op het gebied van informaticatoepassingen. Ten behoeve van de interviews is een aparte maar analoge topiclijst opgesteld om op een systematische wijze gegevens te verzamelen en te ordenen.

In het hierna volgende worden de belangrijkste conclusies samengevat. Daarbij is sprake van een aggregatieprobleem. Het samenvoegen van gegevens van projecten met verschillende doelstellingen, een dwarsdoorsnede door het INSP cluster II.4, daarnaast INSP-LV, NIIO en het overige niet-reguliere onderwijs, vereist een aantal stappen. De eerste is die van evaluatiegegevens per project. Dat is voor een beleidsevaluatie en ons doel nu niet zinvol. Daarom wordt gekozen voor een ordening van conclusies op twee niveaus: per deelprogramma en over het totaal.

4 *Conclusies deelprogramma's*

4.1 *Het reguliere beroepsonderwijs*

Waar het de *algemene doelen* van het INSP cluster II.4 betreft, nl. het creëren van 'human capital', de opvang van de discrepantie tussen aanbod en behoeften op de arbeidsmarkt, de innovatiecapaciteit t.a.v. regulier en niet-regulier onderwijs kan gesteld worden dat de aansluitingsproblematiek in de geanalyseerde projecten niet als centrale doelstelling heeft gefungeerd. Veeleer is een wisselwerking waarneembaar tussen behoeften en initiatieven die vanuit scholen toch al ontstaan waren en de INSP clusterplannen. Veel initiatieven van scholen zijn overgenomen en ingepast. Dat moet dan ook leiden tot de conclusie dat het INSP-project als aanjager heeft gefungeerd of als vliegwiel en daarmee een serie (on)verwachte effecten heeft gesorteerd. Daarop gaan we nu in.

Technology push

De eerste fase van het INSP-programma ook van dit cluster heeft geleid tot een grote apparatuuraanschaf (technology push), zonder dat duidelijk was hoe en op welke wijze die apparatuur zou kunnen worden ingepast in het leerplan en in de instructieprocessen. Uit de software- en implementatieproblémen die later zijn ontstaan blijkt dat. Eerst daarna zijn pogingen ondernomen om leerplan- en leergangontwikkeling (incl. software-aanpassing of -aankoop) en nascholing te regelen. Dat betekent dat niet de leerplanstructuur (de leerinhouden, beroepsopleidingskwalificaties) model heeft gestaan, maar de hardware. Vanuit een leerplantechnisch oogpunt, maar ook vanuit een oogpunt van de inrichting en vormgeving van leerprocessen een verwerpelijke aanpak.

Apparatuur is op grote schaal aangeschaft. Er blijken veel knelpunten op te treden: de levertijden van apparatuur zijn (soms) lang; de apparatuur is vaak niet compatibel; de inrichtingskosten vaak niet voorzien, de onderhoudskosten onvoorzien hoog en het onderhoud matig tot slecht georganiseerd. Overigens menen beleidsvoerders dat er een discrepantie is aan te wijzen tussen dat wat docenten zeggen nodig te hebben – wat vaak wordt toegezegd – en de mate waarin ze erop voorbereid zijn.

Een specifiek probleem in dit verband is de vraag of er gestreefd moet worden naar *standaardisatie*. Er is op grond van de afzonderlijke projectervaringen geen eenduidige conclusie te trekken. Er zijn projecten waarbij op grond van een beredeneerde keuze is gekozen voor een bepaalde configuratie. De ervaringen zijn desondanks wisselend. Er zijn echter ook gevallen waarin apparatuur met succes op de lokale behoeften is afgestemd. Op grond van leerplantechnische, financiële en dissementatie-overwegingen is standaardisatie van besturingssystemen en programmatuur wel aan te bevelen.

De beschikbaarheid van goede software is een probleem. In die gevallen waarin software ontwikkeld moest worden deed men soms een beroep op een educatief softwarehouse als het ECC (voorheen COI). In gevallen waarin men zelf software ontwikkeld heeft inclusief bijbehorende manuals/handboeken rijst het probleem van de auteursrechten en de verspreiding. Scholen hebben daarmee geen ervaring. In enkele gevallen werd samengewerkt met een educatieve uitgever.

Regionale centra

Van ingrijpende aard zijn die projecten waarin getracht is de aansluiting onderwijs arbeidsmarkt via zogenaamde regionale centra verder tot ontwikkeling te brengen, zoals het speerpuntproject MTO en de centra voor kantoor- en winkelpraktijk. Soortgelijke varianten kunnen worden aangetroffen in het NIIO-project en in de landbouw, waar de

proefboerderijen een soortgelijke functie vervullen.

De rationale achter de MTO-centra was dat het opleiden van MTS'ers werktuigbouwkunde en produktietechnieken onder invloed van technologische vernieuwingen onder druk komt te staan. Nieuwe kwalificatie-eisen komen op en deze moeten gerealiseerd worden op ruim 60 scholen. Aangezien voor het trainen van nieuwe vaardigheden kostbare apparatuur vereist is en de middelen beperkt zijn, lag het idee voor de hand om regionale centra uit te rusten met geavanceerde apparatuur die ook gebruikt zou kunnen worden door scholen voor het trainen van vaardigheden op die geavanceerde apparatuur. Vanuit verschillende ministeries (O & W, SOZaWe, EZ) is het initiatief genomen tot de oprichting van regionale centra met de volgende functies:

a. gezamenlijk gebruik van dure apparatuur door verschillende instellingen van beroepsonderwijs;
b. gebruik van geavanceerde apparatuur ter verbetering van de aansluiting van de eindtermen van een opleidingsprogramma van het reguliere onderwijs aan de eisen die aan werknemers in startfuncties worden gesteld door het bedrijfsleven;
c. gebruik van apparatuur voor her-, om- en bijscholing van werknemers en werklozen en nascholing van docenten.

Uit de evaluatie van de projecten blijkt, dat de functie van een regionaal centrum sterk afhankelijk is van de mogelijkheden die aan het MTO worden gegeven in het kader van de SVM-operatie, haar bekostigingsstructuur, de mogelijkheden tot contractonderwijs en cursorisch onderwijs. Verder zijn er tal van politieke ontwikkelingen op provinciaal en lokaal niveau, zoals de regionale besturen arbeidsvoorzieningen (RBA) en het regionaal educatief overleg (REO) waarin overheid, werkgevers en werknemers samenwerken om tot scholingsvoorzieningen te komen. De positie van het regionaal centrum in dat geheel is niet duidelijk.

Indien we de twee in de evaluatie betrokken speerpuntcentra van Almelo en Amsterdam van de in totaal zes bestaande, qua ontwikkeling, opzet, ervaringen en toekomstvisie vergelijken (zie Nijhof et al., 1988, p. 32 e.v.) en beschouwen als exemplarische gevallen voor de overige regionale centra dan trekken we de volgende conclusies:

1. De overheid hinkt t.a.v. de regionale centra op twee gedachten. De regionale centra moeten self supporting zijn, maar de gebruikers van het centrum – de scholen – worden mede door sponsoring (!) goed uitgerust met apparatuur, zodat ze geen gebruik hoeven te maken van een centrum.
2. Een onderscheid in basisvorming door MBO-scholen (basisvaardigheden) en verdieping (van geavanceerde vaardigheden) door regionale centra, een doelstelling die de overheid voor de geest staat, lijkt door bovenstaande ontwikkeling doorkruist te worden.
3. De input van het bedrijfsleven (waar het betreft kennis en middelen) komt aarzelend tot stand. Het is de vraag of het bedrijfsleven voor een structurele basis van centra kan en wil zorgen.
4. Het vermoeden is aanwezig dat het accent zal gaan verschuiven van cursussen t.b.v. het MBO naar cursussen voor het bedrijfsleven. De oorzaak is een veranderde bekostigingsstructuur in het MBO en het feit dat er problemen zijn m.b.t. de samenwerking tussen de scholen. In het kader van de SVM-operatie en de veranderde regelgeving voor het MBO, waardoor contractonderwijs en cursorisch onderwijs mogelijk is, worden de MBO-scholen eerder concurrenten dan partners. Daarmee verliezen de centra een van de beoogde functies.
5. Niet-reguliere opleidingen werken nauwelijks samen met regionale centra. Ze hebben een eigen bekostigingssysteem, eigen doelgroepen, een eigen marketing en een eigen traditie. Ze zijn eerder concurrerend dan coöperatief. De regionale centra lijken tussen wal (MBO-scholen) en schip (irregulier onderwijs) te vallen.

De noodzaak voor MBO-scholen om gebruik te maken van regionale centra is niet aantoonbaar. De apparatuurcomponent van het centrum is op de duur een te geringe factor van betekenis. Het bestaansrecht van centra op termijn wordt daardoor bedreigd. De centra hebben wel een groeiende functie t.a.v. het bedrijfsleven, al ligt de omvang van de dienstverlening ver achter bij de aanvankelijk gemaakte schattingen (Nijhof et al., 1988, p. 33).

Leerplanontwikkeling en nascholing
Een belangrijke taak van veel projecten is het ontwikkelen van voorbeelden van leerplannen, lessen in het licht van een versnelde leerplanontwikkelingsprocedure en een mogelijk langdurig gebruik van lesmaterialen in een verscheidenheid van vormen (o.a. modulen). Er is echter onvoldoende terugkoppeling van de projecten naar het landelijke niveau waardoor de implementatie en afstemming problematisch is. Ook de bepaling van eindtermen voor verschillende vakken wordt onderling niet afgestemd, waardoor afstemmingsproblemen ontstaan met het bedrijfsleven en met HTS'en. Relatieve autonomie binnen projecten zonder outputbewaking en zonder het doen legitimeren van doelen door externe respondenten roept het beeld op van een ongecontroleerde leerplanontwikkelingsstrategie die bottom up werkt in plaats van top down. Gecontroleerde formatieve en summatieve evaluaties vinden ook niet plaats, waardoor het leerproces van ontwikkelaars via trial and error verloopt. Ervaringen zullen zo later moeilijk overdraagbaar zijn naar vervolgprojecten. Er blijkt een onderschatting van de professionele taak van leerplanontwikkeling en een overschatting van de mogelijkheid om buitenlandse leerplanprodukten te adapteren dan wel te reviseren, zowel naar inhoud als naar tempo.

Gelet op deze ervaring kan deze taak beter door professionele instituten worden uitgevoerd of in ieder geval beter worden ondersteund. De roep om snelheid bij leerplanontwikkeling staat in dit geval haaks op de realisatie van de beoogde output. Waar het de nascholing betreft kan worden opgemerkt dat het opmerkelijk is te zien dat eerst na de apparatuuraanschaf de nascholing (veel te laat) wordt georganiseerd, waardoor apparatuur werkloos blijft staan en de invoering in het onderwijs, gecombineerd met leerplanontwikkeling op zich laat wachten.

Organisatie- en implementatieproblemen
Er doemen nogal wat ongelijksoortige problemen op bij de invoering van IT in het onderwijs mede als gevolg van randvoorwaarden die vervuld moeten worden:
- lesuitval bij nascholing van docenten;
- onvoldoende en te late toewijzing van taakuren en/of middelen voor personeel om b.v. software te ontwikkelen (veel liefdewerk);
- bij gebrek aan service van een leverancier onderhouden docenten de apparatuur;
- organisatie van projectonderwijs in relatie tot cursorisch onderwijs geeft problemen te zien;
- gelijktijdig lesgeven in Computer Aided Design (CAD) en in conventionele technieken is niet mogelijk;
- behoud van deskundig personeel is moeilijk: vertrek naar het bedrijfsleven is groot.

De relatie tussen regulier en niet-regulier onderwijs
Binnen de geanalyseerde projecten is de relatie tussen regulier en niet-regulier onderwijs zwak. Als er een vorm van samenwerking is, is deze toevallig en restrictief. Ze beperkt zich tot overname van cursusmateriaal. Een taakafbakening tussen beide soorten onderwijs is er niet.

4.2 *Evaluatie INSP-LV*
In het kader van het INSP-LV in de landbouwsector is vooral in eerste instantie de marktsector gestimuleerd. Tegelijkertijd is gezorgd voor een goede voorlichting en nascholing van docenten uit het landbouwonderwijs. Apparatuurvoorziening en leerplanontwikkeling kwamen op de tweede plaats, vooral omdat men er niet in gelooft dat via het onderwijs technische ontwikkelingen in de marktsector doorgevoerd kunnen worden: onderwijs als aanjager op korte termijn van de economie ziet men dus niet.

De idee van de centralisatie is in het landbouwonderwijs verder doorgevoerd dan in het technische onderwijs. De overheid heeft in het landbouwonderwijs greep op de praktijkcentra, de proefstations, door in de eindtermen de verplichting van een bezoek aan een praktijkcentrum van enkele weken op te nemen. De financiering van deze centra wordt door de overheid geregeld en niet via deelnemende scholen. Zijn binnen een praktijkcentrum opstellingen ontwikkeld (instructieve machines) die ook binnen de scholen gebruikt kunnen worden, dan is het de bedoeling dat dit daadwerkelijk gebeurt. De centra zijn zo toegankelijk voor de markt- en onderwijssector. De relatie tussen overheid, bedrijven en scholen wordt als direct en goed beschouwd. De leerplantrajecten zijn relatief kort.

Waar in het technisch onderwijs vooral is gewerkt met diversificatie in deelterreinen aan de hand van Wagner II en een aanbestedingsprocedure gebaseerd op verzoeken van scholen en dit vooral in de aanvangsfase geleid heeft tot een forse apparatuuraanschaf, daar heeft landbouw voor de weg van de marktverkenning en de leerplanopbouw gekozen in samenhang met een strategie voor onderzoek in proefstations en centra. Deze 'technologische systeemaanpak' is zeer vruchtbaar gebleken: de produkten zijn gereed en ingevoerd in het middelbaar en hoger agrarisch onderwijs en vertoont veel cohesie, maar laat aan het eind van de INSP-LV zien dat in een vervolg wel een technology push noodzakelijk is om de verworven basiskennis informatica toe te kunnen passen. Een belangrijk voordeel van de gekozen weg is dat de gebruikers geïnformeerd zijn en weten wat ze willen. Daardoor is een doelgerichte en efficiënte keuze van hardware mogelijk.

4.3 *Evaluatie van het Nationaal Inhaalprogramma Informatica Opleidingen (NIIO)*

De vraagkant (bedrijfsleven) blijft een lastig te bepalen grootheid. Beschikbare gegevens (Bakkenist et al., 1984) zijn naar het oordeel van respondenten achterhaald. De destijds geformuleerde doelen bleken niet meer actueel. Er is daarom nog wel eens afgeweken van de oorspronkelijke doelstellingen.

De investeringsbereidheid bij bedrijven in nieuwe technologieën (robots, FPA, CAD) is terughoudend. Dat heeft zijn weerslag op de behoefte aan opleidingen daarvoor. Dit gegeven is niet nieuw (Fortuyn, 1985). Wegmann, Chapman en Johnson (1989, p. 22) stellen aan de hand van Amerikaanse literatuur vast dat robots nog tamelijk primitief zijn en het best werken in bedrijven die speciaal voor hen gebouwd zijn. Maar hun eindconclusie is dat de effecten van automatisering en robotisering op de werkgelegenheid moeilijk voorspelbaar zijn gelet op de vele factoren die van invloed zijn op het niveau van werkgelegenheid en werkloosheid. Daarom is elke vorm van scholing en training voor een deel een schot in het duister.

Het NIIO kende, zoals eerder geschetst vijf clusters. Voor de specifieke effecten per cluster verwijzen wij naar Nijhof et al. (1988, p. 20 e.v.).

De algemene effecten zijn dat de aantallen geschoolden niet overdreven hoog te noemen zijn (ongeveer 30% van het verwachte aantal); de kwaliteit van het aanbod is redelijk tot goed te noemen. De respondenten vinden de investeringen in verhouding staan tot de opbrengst, mede gelet op de algemene doelstelling van het programma, nl. het inhalen van een achterstand (Haalbaarheidsstudie Nationaal Inhaal-programma Informatica Opleidingen, april 1984).

Een belangrijke vraag bij deze projecten is of er een uitstraling naar het regulier onderwijs heeft plaatsgevonden. In cluster 3, gericht op de inrichting van Praktijktrainingscentra voor Produktievernieuwing (PTP's) en Automatiseringspraktijkcentra (APC's), is de opzet zo gemaakt dat wordt samengewerkt met de TU Delft op het gebied van CAD/CAM en Robotica. Een uitstraling op landelijk niveau is echter niet vastgesteld.

Effectmetingen vonden niet plaats binnen het NIIO. Het is derhalve niet bekend of er enige doorwerking van nieuwe inzichten plaatsvindt binnen bedrijven. De evaluatie is beperkt gebleven tot het meten van behoefte aan cursussen. Bleef deze op peil dan werd dat als een indicator beschouwd voor marktgerichtheid en kwaliteit van de cursussen.

Nu is het meten van effecten van een cursus naar de werkplek een notoir meetprobleem in het transferonderzoek (zie Cormier & Hagman, 1987; Nijhof & Remmers, 1989). De nu gehanteerde maat is niet sterk te noemen, eerder een teken van onmacht. Een zwakke controle op de output van een toch dure investering.

In het algemeen leed het NIIO onder gebrekkige kwantitatieve en kwalitatieve schattingen. Een permanente updating is noodzakelijk, zoals de General Accounting Office in de Verenigde Staten deze regelmatig pleegt. Inzicht in de kwantitatieve en kwalitatieve verschuivingen is betekenisvol om de opleidingsbehoeften adequaat en snel te kunnen peilen (Kutscher, 1988).

4.4 *Particuliere instellingen: resultaten*

Opleidingsmotieven, deelnemersaantallen en cursuskenmerken

Cursussen worden met name gevolgd vanuit een *carrière*-oogpunt. Het volgen van cursussen varieert van AMBI-modulen voor HBO-

niveau tot eenvoudige trainingen in het gebruik van de PC voor tekstverwerking, spreadsheet, e.d. Dit betekent dat de doelgroepen variëren al naar gelang de vraag. Er is een toename te bespeuren naar HBO-niveau. In het algemeen werd een stijging van het aantal cursisten verwacht, mede door de modulen op AMBI-niveau.

De deelname aan informatica-gerichte dan wel verwante cursussen (peildatum 1987) was groot: ca. 150.000 cursisten. Er is sprake van een forse deelname. Vergelijken we dit cijfer met de deelname van cursisten aan het middelbaar beroepsonderwijs 1985-1986, nl. 280.000 (CBS, 1987, p. 123), dan is er sprake van een aanzienlijke deelname, zij het dat er sprake is van deeltijdcursussen. Waarschijnlijk ligt dit getal nog aanzienlijk hoger wanneer alle mogelijke 'cursussen' en kleine opleidingsinstituten uit de Gele Gids zouden worden meegeteld. Of er sprake is van inhaaleffecten of van een permanente markt is uit deze gegevens niet af te leiden. Sommige organisaties verwachten wel een einde aan de groei en menen een stabilisatie te bespeuren. In ieder geval spelen de bovengenoemde opleidingen een belangrijke rol bij het scholen en trainen van mensen om de aansluiting op, resp. carrière-ontwikkeling en verdere ontplooiing in de arbeidsmarkt te realiseren.

Waar het de aard van de cursussen betreft: bij vier instellingen is sprake van schriftelijke cursussen, bij Teleac is sprake van multimediaal onderwijs en bij NOVI ook van mondeling onderwijs.

In veel gevallen worden cursussen afgesloten met een landelijk erkend examen. Dat geldt in elk geval voor de AMBI-cursussen (Exin, 1987; PBNA, 1987[3]). Instituten geven ook eigen diploma's of certificaten af.

De behoefte aan cursussen bij het bedrijfsleven wordt overigens niet gepeild aan de hand van systematisch (grootschalig) behoeftenonderzoek, maar veeleer door middel van contacten met opleidingsadviseurs, door acquisitie-activiteiten en door het volgen van literatuur, advertenties, en soms door onderzoek.

Planning en Evaluatie

Voorbereiding en uitvoering van onderwijs vindt doorgaans plaats door deskundigen van de instelling zelf plus experts uit het bedrijfsleven én het reguliere onderwijs. Veelal wordt daarbij een plan opgesteld dat conform de planning wordt uitgevoerd door een groep auteurs/docenten/regisseurs/redacteuren. Inhoudelijke en leerplantechnische deskundigheid worden hier gekoppeld. Actualisering van het cursusaanbod vindt gemiddeld eens per drie jaren plaats.

Waar het de cursusevaluatie betreft, deze vindt incidenteel plaats; soms informeel op de werkvloer van het bedrijf, d.m.v. telefonische contacten of d.m.v. huiswerk en in beperkte mate door systematisch onderzoek. De ervaringen van de cursisten worden doorgaans als positief en goed beschreven.

Samenwerking met het reguliere onderwijs

Deze is beperkt en wordt meer bepaald door incidentele behoefte dan door de wil om structureel samen te werken. Voor de cursussen die worden verzorgd ligt naar het oordeel van de respondenten de expertise vaak bij het bedrijfsleven, hoewel men voor de uitvoering van praktijkopdrachten vaak een beroep doet op MTS'en, HTS'en en technische universiteiten i.v.m. de aanwezige technische uitrusting (apparatuur en ruimte).

Conclusie

Het niet-reguliere onderwijs heeft als kenmerk dat het snel kan reageren op ontwikkelingen in het bedrijfsleven. Als het regulier onderwijs bepaalde opleidingen al geeft of gaat geven dan wordt het door het niet-reguliere onderwijs afgestoten: men blijft marktgericht opereren. Dat wil zeggen dat men de wens van de klant, de arbeidsorganisatie, zo veel als mogelijk op maat uitvoert.

Er drukt geen zware wetenschappelijke systematiek op hun werk en er vinden geen grondige evaluatiestudies plaats. Het geheel ademt de sfeer van pragmatiek en snel reageren op een wisselende competitieve markt. Daarin zijn deze bedrijven sterk. Het onderwijsaanbod is niet verrassend nieuw te noemen en loopt zeker niet voorop met geavanceerde programma's.

De huidige *regelgeving* in het reguljer onderwijs wordt als een forse belemmering ervaren om tot samenwerking te komen. Of het reguliere onderwijs een concurrent of een partner wordt zal nader moeten blijken. In de toekomst wordt samenwerking niet uitgesloten als contractonderwijs en cursorisch beroepsonderwijs gerealiseerd worden. In dit

verband wordt ook gerefereerd aan het gebruik van modulen, die tussen regulier en niet-regulier onderwijs uitgewisseld zouden kunnen worden.

Op deze wijze ontstaat een bepaalde verhouding tussen bedrijfsleven, regulier en niet-regulier onderwijs die overigens in het buitenland, m.n. in de Verenigde Staten als normaal wordt beschouwd binnen de zogenaamde community centers. Het reguliere onderwijs verzorgt de permanente scholing van jongeren, t.a.v. basiskennis en -vaardigheden t.b.v. sector, branche of beroepsgroep. Het niet-reguliere (o.a particuliere) onderwijs remedieert, compenseert en differentieert in haar onderwijsaanbod naar marktgerichte groepen met een onmiddellijk te vervullen behoefte. Massaal onderwijs heeft daarbij de voorkeur. In welke mate daardoor specifieke behoeften niet worden afgedekt is niet op voorhand duidelijk (OSA, 1988).

Het bedrijfsleven legitimeert de eisen die voor bepaalde functies gelden mede in het licht van de vraag-aanbod-verhoudingen op de arbeidsmarkt en faciliteert door middel van sponsoring, cursisten, outillage, apparatuur en netwerken de afstemming tussen beide soorten onderwijs. Door interne bedrijfsopleidingen zorgt het voor verder loopbaanbeleid en functiegerichte training.

Door deze vormen van samenwerking gebaseerd op een gemeenschappelijk belang nl. een kwalitatief goed opgeleide manpower, kunnen nieuwe samenwerkingspatronen ontstaan die ook qua kosten-baten-verhoudingen aanmerkelijk gunstiger zijn.

5 *Algemene conclusies*

Verbetering van de aansluiting onderwijs-arbeidsmarkt op het terrein van de informatietechnologie is een van de hoofddoelstellingen van het INSP. Deze aansluitingsproblematiek is met name leidraad geweest.

In deze studie is nagegaan in welke mate de selectie van projecten tegemoet is gekomen aan de algemene doelstelling, namelijk het verbeteren van de relatie onderwijs-arbeidsmarkt. Voorts is nagegaan in hoeverre het niet-reguliere onderwijs wellicht beter in staat is aan te sluiten bij de behoeften op de arbeidsmarkt. Aan deze studie lagen vier vragen ten grondslag:

a. de vraag of human capital wordt gecreëerd door middel van cluster II.4;
b. of de discrepantie tussen onderwijsaanbod en behoeften op de arbeidsmarkt afgenomen is;
c. wat de innovatiecapaciteit van het regulier en niet-regulier beroepsgericht onderwijs is;
d. welke vormen van stimulering noodzakelijk zijn voor de bevordering van informatiekunde resp. informatica.

We beantwoorden bovenstaande vragen aan de hand van de volgende conclusies.

Human Capital

De analyse heeft laten zien dat de aansluitingsproblematiek in de meeste projecten niet als centrale operationele doelstelling heeft gefungeerd. Systematische behoeften- en beroepenanalyses en analyses van opleidingsprofielen hebben nauwelijks plaatsgevonden. Leerplanontwikkeling is daarop niet gebaseerd geweest.

Nieuwe technologieën oefenen invloed uit op de be- en verwerkingsprocessen in arbeidsorganisaties en vergen nieuwe kennis, vaardigheden en attitudes van medewerkers.

Nieuwe beroeps- en functieprofielen komen op. In een aantal gevallen zijn leerplannen ontwikkeld op basis van beroepsprofielen. Welke richting de kwalificaties precies zullen ingaan is nog niet duidelijk, daarvoor is de ontwikkeling van IT nog te weinig gepenetreerd en gestabiliseerd. Wel is een trend waarneembaar naar generale vaardigheden, naar basis- of fundamentele vaardigheden (Nijhof & Mulder, 1986). Deze vaardigheden echter variëren per vakgebied en vertonen ook differentiaties naar functies (Hesse & Nijhof, 1988). Zo hanteert het NGI (1986) 17 algemene vaardigheden voor groepen informatici.

Of kwalificaties na training direct leiden tot de beoogde functies wordt ontkracht door een uitstroomonderzoek onder MTS'ers (Kanselaar et al., 1987), waaruit blijkt dat 50% van de uitstromers van het MTO niet in specifieke MTO-functies terecht komt. Nijhof en Mulder (1986) vonden soortgelijke effecten bij LEAO-schoolverlaters. Het cluster II.4 heeft kortom in beperkte mate bijgedragen tot de ontwikkeling van kwalificaties.

Het NIIO-programma sloot directer aan bij behoeften aan bepaalde technische vaar-

digheden zoals op het terrein van CAD/CAM en robotica. De effecten daarvan in bedrijven zijn echter niet vastgesteld. De grote vraag naar cursussen in het particuliere onderwijs kan als een maat worden beschouwd van het belang dat men hecht aan bepaalde kennis en vaardigheden. Ook hier echter ontbreekt een toets op de transfer van het geleerde naar de werkplek.

Binnen de projecten is wel een duidelijk begin gemaakt met de ontwikkeling van programma's, zijn cursussen georganiseerd en is de aanzet gegeven tot kwalificatiebevordering. Dat is echter het meest systematisch gebeurd bij Landbouw. Daar is sprake van een grote mate van planmatigheid, interne samenhang en doelgerichtheid. De echte invoering zal echter moeten plaatsvinden via het programma PRINT dat begin 1989 van start is gegaan.

Vraag en aanbod verhoudingen

Als de vraag naar de kwalitatieve aansluiting moeilijk te beantwoorden is, dan is het de vraag of kwantitatieve schattingen mogelijk zijn. Het bedrijfsleven tast hier zelf ook in het duister, omdat zij niet exact weet welke richting de arbeidsorganisaties opgaan, welke eisen nodig zullen zijn. Wel treedt een trend tot verdringing op, waardoor hoger geschoolden lager geschoolden wegdrukken uit hun functie (zie b.v. OSA, 1988). De kwantitatieve studies van Bakkenist, Spits en Co (1984) die de basis vormden voor het NIIO-programma zijn modelmatige schattingen en bleken snel achterhaald (zie ook Van Hoof, 1986). Fortuyn heeft in 1985 schattingen laten zien op grond van zeer voorzichtige analyses van de ontwikkeling van Flexibele Produktie Automatisering en robotisering in de Nederlandse situatie. In de USA heeft men zijn schattingen moeten bijstellen in neerwaartse richting (Levin, 1984). Wegmann et al. (1989) schetsen een ambivalente situatie met als netto effect een mogelijk verlies van banen op termijn. De concurrentie wordt harder en het opleidingsniveau dat gevraagd wordt voorlopig hoger (Rumberger, 1988, p. 87). Rumberger wijst echter ook op de handel en de dienstverlening waar de-skilling optreedt en de lonen lager zijn dan in andere delen van de economie, '(...) but (...) many workers may be employed in industries and occupations that require no more than a *moderate level of basic skills* even if they actually use quite sophisticated technical products to perform their jobs" (Rumberger, 1987, p. 86).

Voor de ontwikkeling van opleidingen is de analyse van beroepen noodzakelijk, maar niet als de enige determinant voor leerplanontwikkeling voor het beroepsonderwijs. Evenzeer zijn de analyses van trends in kennisontwikkeling noodzakelijk. Door een meervoudige analyse van beroepen en door analyse van trends en de transformatie daarvan in termen van inhoud en instructie, voorkomt men een te naïeve vertaalslag die conserverend werkt (Brandsma & Nijhof, 1989). Periodieke analyses van beroepsontwikkelingen, kennisdomeinen en organisatie-ontwikkeling zijn nodig om opleidingen up to date te houden (Carlisle, 1985; Nijhof, 1986).

Verondersteld mag worden dat in de periode 1984-1988 de discrepantie tussen vraag en aanbod op de arbeidsmarkt is verminderd, eenvoudig omdat er een groter aanbod is gekomen, mede door het INSP, dat er voor die tijd niet was en door initiatieven van mensen zelf waardoor de vraag tijdelijk is toegenomen. Of daardoor ook nieuwe behoeften zijn gecreëerd is echter niet duidelijk. Het programma werkte als een vliegwiel met deels ongecontroleerde effecten.

Innovatiecapaciteit

Op de vraag of niet-regulier onderwijs flexibeler kan aansluiten op de behoeften van de arbeidsmarkt, is het antwoord positief, maar de aansluitingsproblematiek wordt hier evenmin systematisch aangepakt.

De ervaringen uit het Nationaal Inhaalprogramma laten zien dat het overigens niet vanzelfsprekend is dat niet-regulier onderwijs het zonder meer beter doet. Het particuliere onderwijs (een zeer heterogeen conglomeraat van opleidingen en cursussen) speelt echter wel soepel in op de behoeften van de arbeidsmarkt, is minder gebonden aan regelgeving, verroostering, bevoegdheden en exameneisen. Het kan inspelen op de markt en moet berekende risico's nemen. Het mikt qua cursusaanbod op massale groepen, en op risicoloos onderwijs.

Naar de aard van zijn organisatie is regulier onderwijs strakker aan regels gebonden, dient het te letten op zijn instroom, de exameneisen

te realiseren, en is het verder gebonden aan andere wettelijke voorschriften terzake inrichting en vormgeving.

Uit het feit dat veel projecten in Cluster II.4 voortgekomen zijn uit initiatieven van scholen zelf, dat deze soms sponsoring door het bedrijfsleven hebben weten te organiseren, kan afgeleid worden dat ze wel degelijk over innovatiecapaciteit beschikken. De randvoorwaarden zijn echter vaak ongunstig en soms ontbreekt de managementcapaciteit om een project te 'runnen'.

Naarmate het reguliere onderwijs de mogelijkheid krijgt contractonderwijs uit te voeren kan het flexibeler opereren, maar loopt het de kans om evenals het particuliere onderwijs conserverend te opereren (Nijhof, 1989).

De bevordering van informatiekunde

De uitvoering van de projecten is sterk bepaald geweest door de apparatuurcomponent. Curriculumontwikkeling, softwarevoorziening en nascholing liepen duidelijk achter bij de apparatuurverstrekking. Dit heeft niet alleen geleid tot een sterke onderbenutting van de apparatuur, maar ook tot een verkeerde werking. De kwaliteit van de apparatuur dicteerde de software-ontwikkeling of -aanschaf, of werd in het geheel niet of voor ander doeleinden gebruikt.

Van het grootste belang voor de ontwikkeling van een vakgebied informatiekunde is het bepalen van de harde kern van dat vakgebied en van al zijn toepassingen in de beroepsgerichte context. Veel problemen in de sfeer van de leerplanontwikkeling zijn terug te voeren tot onhelderheid over de structuur van een vak, de kernonderwerpen, de eindtermen en de beheersingsniveaus (zie ook WRR, 1986). De gehele discussie over het core curriculum, dat is niet hetzelfde als basisvorming (!), is daarop gebaseerd. Eerst komt de vraag naar het wat (waarom en inhoud) en daarna de hoevraag (instructie). Deze essentiële kwestie, al zo oud zolang er leerplannen worden gemaakt, wordt vaak om wille van opportunistische motieven en overwegingen opzij gezet. Men kan er echter niet om heen. De vraag naar de stimulering van de informatiekunde en zijn toepassingen zal dus in de eerste plaats moeten gaan over de vraag wat de inhoud en de concepten van het kennis- en toepassingsgebied informatiekunde zijn. Daarna zijn experimentele prototypische ontwikkeling en beproeving van programma's noodzakelijk.

Exploratief ontwikkelingsonderzoek is noodzakelijk om uit te maken onder welke voorwaarden en met welke effecten mensen die kennis integreren en benutten: bijvoorbeeld d.m.v. modulaire systemen (Nijhof & Van den Berg, 1987), door middel van simulaties (Pieters & Treep, 1989), of door toepassing van expertsystemen. Regionale centra kunnen daarbij als proefstation fungeren evenals scholen.

Nader onderzoek naar ontwikkelingsmethodieken voor beroepsgerichte onderwijsprogramma's en de implementatie daarvan in het onderwijs is bovendien geboden.

Het laat zich aanzien dat in PRINT een aanzet wordt gemaakt tot een serieuze prototypische ontwikkeling van programma's in nauwe relatie tot nascholing, implementatie en apparatuurconfiguraties. Er resteert één grote zorg: het evidente gebrek aan geplande output gerichte evaluatiestudies.

Literatuur

Bakkenist, Spits & Co, *Haalbaarheidsstudie Naar een Nationaal Inhaalprogramma Informatica Opleidingen.* Den Haag: 1984.

Brandsma, J. & W.J. Nijhof, *Vergelijkende evaluatie van methoden voor de ontwikkeling van beroeps(opleidings)profielen.* Paper gepresenteerd op de Onderwijs-researchdagen 1989 te Leiden, 24 en 25 mei 1989.

Carlisle, K.E. & J.W. Arwady, *Analyzing Jobs and tasks.* Englewood Cliff, NJ: Educational Technology Publications, 1985.

CBS, *Statistisch Zakboek.* Den Haag: Staatsuitgeverij, 1987.

COB/SER, *Nieuwe technologieën: veranderingen in bedrijf en onderwijs.* Den Haag: 1985.

Cormier, S.M. & J.D. Hagman, *Transfer of Learning.* Contemporary Research and Applications. New York: Academic Press, 1987.

Exin, *AMBI 88 Voorstel tot vernieuwing van de AMBI-examens.* Amsterdam: 1986.

Fortuyn, W.S.P., *Om de toekomst van de werkgelegenheid.* Kampen: 1985.

Hesse, C.G. & W.J. Nijhof, *DACUM in de Boekenbranche.* Technologische ontwikkelingen en de implicaties voor de opleiding tot boekhandelaar en uitgever. Amsterdam/Enschede: Stichting Speurwerk Betreffende het Boek/Vakgroep Curriculumtechnologie, Universiteit Twente: 1988.

Hoof, J. J. van, *Aansluiting tussen onderwijs en Arbeid: Recente Ontwikkelingen en Uitgangspunten voor Beleid.* OSA-werkdocument nr. W 28. Den Haag: 1986.

IBM, *Boer en tuinder op weg naar managementsystemen.* Eindhoven: 1988.

Informaticastimulering, *Voortgangsrapportage.* Den Haag: Staatsuitgeverij, 1987.

Kanselaar, G., J. de Bruijn, J. Cadot, F. Dehnling, A. Groot & H. Maas, *MTO-uitstroom project: Onderzoek naar de aansluiting van het MTS-onderwijs naar de Arbeidssituatie.* (Conceptrapport + bijlagen) Utrecht: ISOR, Rijksuniversiteit, 1987.

Kutscher, R., An Overview of the Year 2000. *Occupational Outlook Quarterly,* 1988, 32, 2-9.

Levin, H. M., *Education and Jobs in a Technological World.* The National Center for Research in Vocational Education. Information Series, no. 265. The Ohio State University: 1984.

Mulder, M., J. Akkerman & N. Bentvelsen, *Bedrijfsopleidingen in Nederland.* Den Haag: Instituut voor Onderzoek van het Onderwijs, 1988.

Nederlands Genootschap voor Informatica, *Functies in de Informatica. Typering, plaats, functievereisten, loopbaanmogelijkheden.* Amsterdam: NGI, 1986.

Nijhof, W. J., *Van Beroepsprofielen naar Curriculum-profielen.* Enschede: Onderzoekscentrum Toegepaste Onderwijskunde, Universiteit Twente, 1986.

Nijhof, W. J., *Privatiseren: Enkele onderwijskundige gevolgen van een economische aanpak.* Voordracht gehouden ter gelegenheid van het symposium Privatiseren in het Onderwijs bij het 10 jarig bestaan van de Faculteit der Toegepaste Onderwijskunde op 16 juni 1989. Enschede: Universiteit Twente, 1989.

Nijhof, W. J. & E. van den Berg, *Moduleren Technische Informatica in het MTO* (onderzoeksvoorstel t.b.v. SVO). Enschede: Onderzoekscentrum Toegepaste Onderwijskunde, Universiteit Twente, 1987.

Nijhof, W. J. & M. Mulder, *Basisvaardigheden in het Beroepsonderwijs.* Den Haag: SVO, 1986.

Nijhof, W. J. & J. L. M. Remmers, *Basisvaardigheden nader bekeken.* Een literatuurstudie naar de inhoud van het begrip basisvaardigheden en een poging tot operationalisering. Enschede: Faculteit Toegepaste Onderwijskunde, Universiteit Twente, 1989.

Nijhof, W. J., Reijmerink, C. H. L. & J. A Thijsen, *Informatietechnologie in het regulier en niet-regulier beroepsgerichte onderwijs.* Den Haag: Staatsuitgeverij, 1988.

OSA, *Arbeidsmarktperspectieven.* OSA-rapport nr. 5. Den Haag: Staatsuitgeverij, 1988.

PBNA, *Poly-automatiseringszakboekje* (red. Th. M. A. Bemelmans, J. A. van der Pool & N. J. M. Zwaneveld). Arnhem: PBNA, 1987[3].

Pieters, J. & M. Mulder (red.), *Praktische problemen in het Beroepsonderwijs.* Enschede: OCTO, Universiteit Twente, 1988.

Pieters, J. & J. Treep, *Simulaties in het beroepsonderwijs.* Enschede: Faculteit der Toegepaste Onderwijskunde, Universiteit Twente, 1988.

Produktiviteitscentrum Metaalnijverheid, *Functieveranderingen als gevolg van technologische ontwikkelingen.* De Bilt: 1985.

PSOI-reeks nr. 20, *Beleidskader Informatietechnologie in het beroepsonderwijs v.o. 1987 en 1988.* Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1987a.

PSOI-reeks nr. 27, *Overzicht per onderwijscluster van goedgekeurde INSP-projecten.* Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1987b.

PSOI-reeks nr. 34, *Beleidsnotitie OPSTAP.* Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1988.

Reijmerink, C. H. L., *Regionale voorzieningen in de werktuigbouwkunde: een structurele oplossing?* Enschede: OCTO, Universiteit Twente, 1987.

Rumberger, R. W., The potential impact of technology on the skill requirements of future jobs in the US. In: G. Burk & R. W. Rumberger (Eds.), *The future impact of technology on work and education* (pp. 74-95) London: Falmer Press, 1987.

Scheerens, J. (red.); *Evaluatie: om de kwaliteit van het onderwijs.* Lisse: Swets & Zeitlinger, 1989.

Scheerens, J., De evaluatie van het onderwijsdeel van het Informatica Stimuleringsplan; evalueerbaarheidsanalyse en formatieve evaluatie. *Pedagogische Studiën,* 1990, *67.*

SOM, *Jaarverslag Ontwikkelingsprojecten 1987; Jaarplan Ontwikkelingsprojecten 1988.*

Spenner, K., The Upgrading and Downgrading of Occupations: Issues, Evidence and Implications for Education. *Review of Educational Research,* 1985, *55,* 135-154.

Twijnstra Gudde NV, *Vooronderzoek Haalbaarheid Vormgeving Technische Praktijkcentra.* Deventer: 1987.

WRR, *Basisvorming in het Onderwijs.* Den Haag: Staatsuitgeverij, 1986.

Curriculum vitae

W. J. Nijhof (1941) is sinds 1981 hoogleraar onderwijskunde, Faculteit der Toegepaste Onderwijskunde, Universiteit Twente. Hij houdt zich o.a. bezig met de onderwijskundige aspecten van het kwalificatievraagstuk in het bijzonder met betrekking tot bedrijfsopleidingen, beroeps-(opleidings)profielen en basisvaardigheden.

Adres: Vakgroep Curriculumtechnologie, Postbus 217, 7500 AE Enschede

Manuscript aanvaard 6-12-'89

Summary

Nijhof, W. J. 'Fly wheel effects of a stimulation program on information technology.' *Pedagogische Studiën*, 1990, *67*, 100-115.

The introduction of Information Technology in The Netherlands has been stimulated by a Program of the Dutch Government called INSP. Part of this program was one oriented at vocational education and training in business and industry. Goals of the program were
- to foster the impact of new information technologies (NIT) into courses and programs on the basis of new qualification and job profiles,
- to test configurations of organizational cooperation between the formal education system and the non formal training system.

The overall effect of the program is a technology push and in service training after the facts. School based curriculum development (software and courseware) has led to products not being tested and therefore without nationwide dissemenation. Some experiments with local and regional centres for training in technical skills (CAD/CAM robotics) were succesful in relation to business and industry, but less to the formal education system. A model of system engineering in agriculture related to the introduction of applications of NIT in curricula and experimental farms proved to be very succesful. The main results will be implemented by means of a new program PRINT. Hopefully, an output oriented evaluation study will follow this program.

Mét of zonder basisvorming?
Evaluatie in het voortgezet onderwijs

J. L. PESCHAR
Vakgroep Sociologie, Rijksuniversiteit Groningen

Samenvatting

In 1989 zou de basisvorming in het voortgezet onderwijs worden ingevoerd. De evaluatie van de voorgenomen doelstellingen bij de invoering is reeds in een vroeg stadium gepland. In dit artikel wordt de achtergrond van de evaluatie geschetst. De doelstellingen van de basisvorming zijn vertaald in evaluatievragen, die in een uitgebreid evaluatieplan zijn opgenomen. Hoewel er op dat moment van invoering van de basisvorming nog geen sprake is, zijn de eerste stappen van evaluatieonderzoek inmiddels gezet. Daarbij worden drie leerlingencohorten op relevante momenten getest en in de schoolloopbaan gevolgd, zowel in het huidige onderwijs als in scholen met basisvorming. De verschillende voorstudies en deelonderzoeken worden kort besproken, alsmede enkele consequenties van de evaluatie. Ondanks de uiterst trage politieke besluitvorming over de basisvorming is in twee opzichten winst geboekt: de erkenning dat evaluatie een essentieel onderdeel is van grootschalige innovaties en de start van een grootschalig onderzoeksprogramma naar de effectiviteit van schoolkenmerken in het voortgezet onderwijs.

1 *Inleiding*

Nadat in mei 1989 het kabinet Lubbers II demissionair was geworden, is lange tijd het lot van de in 1985 voorgestelde basisvorming onduidelijk gebleven. Na een periode van onzekerheid werd eerder al in januari 1989 bekend gemaakt dat de invoering van deze lang verwachte hervorming van het voortgezet onderwijs naar 1990 wordt doorgeschoven. Tot op dàt moment hadden het parlement, het onderwijsveld, de ouders en evaluatie-onderzoekers steeds in de onzekerheid verkeerd of het inderdaad zou lukken om de basisvorming nog in 1989 te introduceren. Met het regeerakkoord van het kabinet Lubbers III is komen vast te staan, dat het inderdaad de bedoeling blijft de basisvorming definitief in te voeren. We spreken dan op zijn vroegst van 1991.

2 *Achtergrond*

In vergelijking met andere landen in Europa werd in Nederland de integratie in de eerste fase van het voortgezet onderwijs relatief laat overwogen. De eerste echte onderwijsverandering vindt pas in de jaren zestig plaats: de invoering van de Mammoetwet (1968). De onbedoelde en nadelige effecten van het nieuwe systeem van brugklassen voor algemeen vormend en beroepsonderwijs werden reeds snel onderkend en al in het begin van de jaren zeventig werden andere ideeën voor een comprehensive-achtig schooltype in Nederland ontwikkeld. Door minister Van Veen werd in 1972 (Nota over het onderwijsbeleid) de gedachte van geïntegreerd voortgezet onderwijs van 12- tot 15-jarigen als uitgangspunt genomen. In 1973 worden plannen voor de experimentele Middenschool aangekondigd. Een verdere uitwerking van deze plannen wordt in de Contourennota van 1975 gepresenteerd. In 1976 gaan de eerste experimenten van start, waarmee het mogelijk zou moeten zijn om na te gaan of een schooltype kan worden gerealiseerd waarin grote groepen leerlingen gezamelijk het onderwijs volgen, ongeacht aanleg, milieu of prestatie.

De loop der geschiedenis heeft echter anders beslist. De experimenten met de Middenschool werden niet verder uitgebouwd. Tegelijkertijd is ook de evaluatie van de Middenschoolexperimenten nauwelijks ter hand genomen. Politiek gesproken zou in het voortgezet basisonderwijs, dat in 1985 werd voorgesteld, een deel van de doelstellingen van de Middenschool kunnen worden opge-

Pedagogische Studiën 1990 (67) 116-126

nomen. Een nieuw element bij de voorstellen voor het voortgezet basisonderwijs was, dat de invoering van dit schooltype *afhankelijk* zou worden gemaakt van de bevindingen van het begeleidend evaluatieonderzoek. Daartoe wordt een Coördinatie Commissie Evaluatie Voortgezet Onderwijs (CCE) door de minister ingesteld. Met name deze commissie heeft een belangrijke stap gezet in de richting van beleids-*on*afhankelijke evaluaties met een vergelijkend karakter. Hierop zal verderop nog worden ingegaan.

Inmiddels bleken echter de voorstellen over het voortgezet basisonderwijs door de politieke realiteit te zijn achterhaald. In 1986 brengt de Wetenschappelijke Raad voor het Regeringsbeleid een uitgebreid advies uit over basisvorming in het onderwijs. Hiermee wordt de tot dan toe gekozen koers van onderwijsvernieuwing door *structuur*veranderingen definitief verlegd naar de verandering van de *inhoud van het onderwijs* (WRR 1986, hoofdstuk 4). Het voorstel van de WRR over de basisvorming is zeer snel gemeengoed geworden en fungeerde als uitgangspunt voor de kabinetsformatie van 1986. Zelfs nog aan het eind van hetzelfde jaar verschijnt een conceptwetsvoorstel waarmee de basisvorming zou kunnen worden gerealiseerd. Essentieel aan dit voorstel is bovendien dat er een *expliciete* paragraaf aan de evaluatie van het invoeringsproces is toegevoegd[1].

Hoewel in dit wetsvoorstel nog niet genoemd, lag het voor de hand om de CCE met deze evaluatie te belasten. Die had inmiddels gekozen voor een bijna *experimentele* onderzoeksaanpak voor de evaluatie van het VBaO. Daarbij zouden de experimentele scholen op een groot aantal relevante kenmerken worden vergeleken met het bestaande onderwijs (CCE 1985). Tegelijkertijd werden *van te voren criteria* gekozen die bij de beoordeling van uitkomsten zouden worden gehanteerd. Met dit eerste evaluatieplan werd in feite de standaard gezet voor het verdere evaluatie-onderzoek op dit terrein. Betrekkelijk uniek is ook dat na de aanbieding van het advies door het ministerie geen verdere wijzigingen op dit model zijn voorgesteld. Hiermee was overigens wel een merkwaardige situatie ontstaan, namelijk dat er wel een onderzoeksplan bestond voor een onderwijstype dat nooit gerealiseerd zou gaan worden (zie ook Kremers, 1985).

Bij de indiening van het definitieve wetsvoorstel in december 1987 bleek dat de uitgangspunten voor de invoering van de basisvorming onveranderd waren en de condities voor het evaluatieonderzoek evenmin zouden veranderen. De CCE zou de evaluatie gedurende de tienjarige invoeringsperiode gaan coördineren waarbij in het Vierde Advies aan de minister (februari 1987) de eerste voorstellen zijn gedaan. Er is een onafhankelijke evaluatie bepleit, die een zekere afstand tot de politiek bewaart. Nadat de randcondities waren vastgesteld, kon met de blauwdruk van de evaluatie worden begonnen. Deze is recent beschikbaar gekomen als integraal onderdeel van het Zesde Advies van de CCE aan het ministerie (1988)[2].

3 *Het evaluatiemodel*

3.1 *Referentiegroepen en vergelijkende vraagstellingen*

Het grootschalige evaluatie-onderzoek in Nederland heeft een roemruchte geschiedenis. De verschillende innovaties in jaren zeventig (Open School, Middenschool en innovatieprojecten) werden vanuit evaluatie-oogpunt kritisch besproken door Scheerens (1983, 1987). Het belangrijkste probleem hierbij was vooral dat evaluatievragen niet waren gespecificeerd en een vergelijking ten opzichte van *referentiegroepen* niet werd gerealiseerd. In dit opzicht is er dus sprake van een duidelijke koerswending: reeds vóórdat de innovatie wordt ingevoerd, is er sprake van een evaluatiekader.

Een tweede belangrijk winstpunt is, dat er inderdaad een evaluatie *mogelijk* is. Immers de invoering van de basisvorming wordt gerealiseerd in een periode van tien jaar, waarbij de situatie gaat ontstaan dat een aantal scholen is overgegaan op een model voor basisvorming en andere nog niet. Dit opent in principe een mogelijkheid voor de vergelijking tussen twee verschillende typen scholen.

Een tweetal typen evaluatievragen is hierbij relevant. In de eerste plaats kan er sprake zijn van een vergelijking *tussen* scholen op een bepaald moment, die wèl of niet basisvorming hebben ingevoerd. In de tweede plaats kan – omdat er immers gedurende een tien-jaren periode wordt ingevoerd – worden nagegaan of er eventuele verschillen tussen de schooltypen *in de loop der tijd* groter dan wel kleiner wor-

den. In het evaluatieplan is aan deze twee typen vraagstelling de prioriteit gegeven. Natuurlijk is het ook mogelijk om te besluiten *binnen* scholen met basisvorming te gaan onderzoeken, maar het uitgangspunt van een referentiegroep wordt daarmee niet gehonoreerd.

3.2 *Meetinstrumenten*

Een cruciaal punt in *alle* vergelijkende onderzoeken – en dus ook bij evaluatieonderzoek tussen twee verschillende onderwijssystemen – is de beschikbaarheid van een gemeenschappelijk meetinstrument. Het is de bedoeling dat de basisvorming wordt afgesloten met het afleggen van toetsen op een veertiental terreinen. Hiermee is tegelijkertijd duidelijk dat deze toetsen geen geschikt criterium voor de evaluatie zijn: ze zijn immers niet beschikbaar voor het 'oude' onderwijssysteem. Daarom moest er voor de evaluatie een apart criterium worden ontwikkeld; de toetsen hiervoor zijn door het CITO uitgewerkt. Na uitvoerige discussie is de keus gevallen op een viertal toetsen die betrekking hebben op Nederlands, Engels, Biologie en Wiskunde. Deze toetsen worden gestandaardiseerd aan het einde van het derde leerjaar afgenomen. Terzijde kan nog worden vermeld dat ook om een andere reden de toetsen basisvorming geen aantrekkelijk criterium zijn. Immers scholen zelf hebben de mogelijkheid om te kiezen *op welk moment* zij in het tweede, dan wel derde leerjaar de toetsen van de basisvorming zullen afnemen. Dit betekent voor onderzoek dat op ongelijke momenten wordt gemeten: zodoende worden reeds verschillen tussen scholen 'ingebakken'.

4 *Vraagstellingen van de evaluatie*

4.1 *De wetgeving basisvorming en het achterliggende evaluatiemodel*

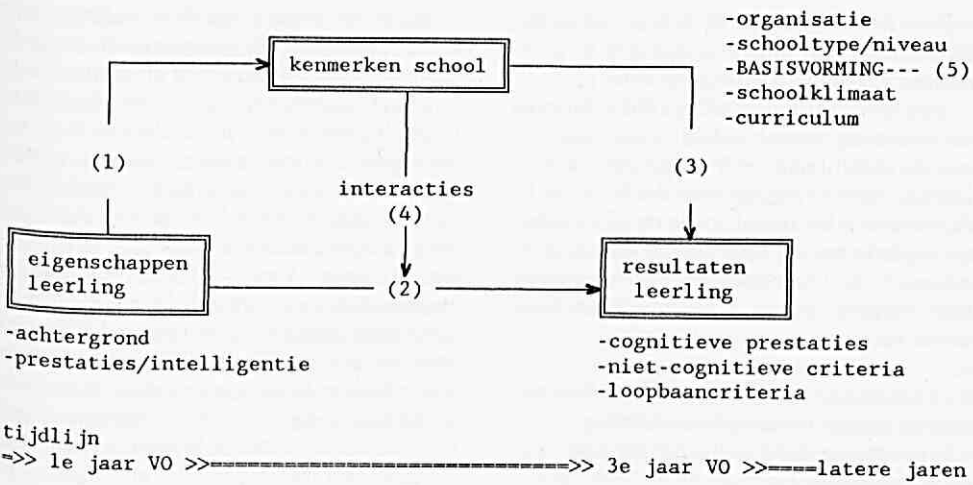
De inhoud en vormgeving van de voorgenomen basisvorming mogen zo langzamerhand bekend worden verondersteld (zie daarvoor het *Wetsontwerp Basisvorming* en het *Invoeringsplan Basisvorming*). Voor de evaluatie is, zoals gezegd, vooral van belang dat een periode van tien jaar kan worden gebruikt om de basisvorming geleidelijk aan in te voeren. Een belangrijk *nadeel* hiervan is dat scholen zich zelf kunnen aanmelden voor evaluatie, zodat er met name in de vergelijkende evaluatie sprake zou kunnen zijn van een selectie-effect. Hierop moet in de analyses terdege worden gecontroleerd.

In de tekst van de Memorie van Toelichting worden slechts globale aandachtspunten van evaluatie gevraagd. In een eerste lezing heeft de CCE (Vierde Advies, februari 1987) uit deze Memorie van Toelichting een aantal evaluatievragen gedestilleerd. Om na te kunnen gaan of dergelijke vragen voldoende ruimte boden voor de beoogde evaluatie, is daarna een afzonderlijke analytische evaluatie gehouden. In deze *vooraf-evaluatie* is door een drietal auteurs uitvoerig nagegaan onder welke condities de evaluatie zou moeten gaan starten en wat eventuele *bottlenecks* in de uitvoering zouden kunnen voorkomen (zie Beem, 1988). Uit de analytische evaluatie blijkt duidelijk dat er sprake is van verschillende niveaus waarop de evaluatie zich afspeelt: dat van het individu, de klas, de school, het onderwijstype, het nationale onderwijssysteem. Dit vereist dat van te voren de onderzoeksvragen goed zijn gespecificeerd op de relevante niveaus. Om echter de analyse op de verschillende niveaus te kunnen uitvoeren, is de verzameling van gegevens over *leerlingen* een eerste vereiste: er kan immers altijd worden geaggregeerd naar een hoger niveau, maar niet naar een lager niveau. Dit leidt tot de definitieve keuze van het onderzoeksmodel.

Op basis van de in paragraaf 3 genoemde eisen kan nu het achterliggende onderzoeksmodel worden vastgelegd en gekoppeld aan de inhoudelijke vragen van de evaluatie. In Figuur 1 is weergegeven welke relaties worden verondersteld tussen de verschillende blokken variabelen. De achterliggende gedachte is dat de kenmerken van de leerlingen – met name de prestaties – door de factoren schoolniveau in gunstige zin worden beïnvloed. Vinden we inderdaad dat de scholen met een basisvormingspakket gekenmerkt worden door een verhoging van de kwaliteit van het onderwijs, onder constanthouding van de leerlingkenmerken?

Dit houdt in dat we in eerste instantie geïnteresseerd zijn in de relatie tussen schoolkenmerken en resultaten (relatie 3), onder gelijktijdige uitschakeling van andere invloeden: de eigenschappen van de leerling.

Gelijktijdig aan deze verzameling van gegevens voor de evaluatie op *individueel (mi-*



Figuur 1 *Het achterliggende basismodel van de evaluatie*

cro) niveau, kan ook op *macro niveau* een evaluatie worden opgezet (zie Dronkers, 1988). Daartoe kunnen bestaande en nieuw te ontwikkelen tijdreeksen relevante informatie opleveren. Er moeten dan aparte indicatoren voor ons onderwijsstelsel worden geconstrueerd, zodat er sprake is van een situatie waarin tijdreeksen van onderwijsindicatoren breuklijnen gaan vertonen op het moment dat de basisvorming zou zijn ingevoerd. Het Centraal Bureau voor de Statistiek heeft in dit opzicht een voorstudie verricht (Schouten, 1987; CBS, 1988), waarin een groot aantal indicatoren is geformuleerd en waarvoor de gegevensbewerking inmiddels in gang is gezet.

4.2 *De evaluatievragen*

In het verlengde van de doeleinden van basisvorming is een zestal typen evaluatievragen met verwachtingen (hypothesen) geformuleerd. Elke vraag is vervolgens op te splitsen in specificaties voor schooltypen, prestatiegroepen, leerlingcategorieën, en dergelijke. Voor de overzichtelijkheid wordt hier volstaan met de *zes typen hoofdvragen A t/m F.*

A: *Evaluatievragen over de verhoging van het onderwijspeil*

A1: Zijn bij leerlingen met vergelijkbare kenmerken de prestaties in de basisvorming hoger dan in het huidige onderwijs?

A2: Worden geconstateerde verschillen in prestaties tussen basisvorming en het huidige onderwijs bij leerlingen met vergelijkbare kenmerken in de loop van de tijd groter en in welke richting?

In deze vragen wordt een tweetal gezichtspunten naar voren gebracht. In de eerste plaats is de veronderstelling dat, gegeven gelijke instroomkenmerken, de prestaties in de basisvorming hoger zullen zijn dan die in het bestaande onderwijs. Het onderwijspeil wordt geoperationaliseerd in termen van prestatie- en non-cognitieve toetsen, voor zover in de school af te nemen. Daarna kan alleen in termen van de positie in de schoolloopbaan worden geëvalueerd. Er wordt vanuit gegaan dat, naarmate er meer scholen in de loop van de tijd hebben kunnen inspelen op de basisvorming, de verschillen met het bestaande systeem gaandeweg groter zullen zijn. Deze uitgangspunten moeten vooral gezien worden als *hypothesen* waartegen de analyses later zullen kunnen worden afgezet. Feitelijk zijn het ook de doelstellingen voor invoering van de basisvorming.

B: *Vragen over de versterking van de gemeenschappelijke culturele basis*

B0: Is het inderdaad zo dat het leerpakket in de basisvorming breder en homogener is samengesteld voor grotere groepen leerlingen dan in het huidige onderwijs?

B1: Hebben leerlingen in de basisvorming in vergelijking met soortgelijke leerlingen in het huidige onderwijs inderdaad een bredere culturele basis?

B2: Wordt de gemeenschappelijke culturele

basis in de basisvorming in de loop van de tijd groter, in vergelijking met soortgelijke groepen leerlingen in het huidige onderwijs?

Een eerste veronderstelling (B0) is dat in de basisvorming een inhoudelijk ander curriculum zal gaan functioneren. Door inhoudsanalyse kan worden nagegaan of dat het geval is. Pas daarna is het zinvol om de verdere evaluatievragen te stellen. Overigens is de wet en de memorie van toelichting betrekkelijk duister over datgene wat er onder culturele basis wordt verstaan.

C: *Vragen met betrekking tot de technologische en sociaal-culturele ontwikkeling*
C0: Is het inderdaad zo dat het curriculum in de basisvorming meer aandacht besteedt aan technologische en sociaal-culturele ontwikkeling dan in het huidige onderwijs?
C1: Is de arbeidsmarktpositie van leerlingen na de basisvorming beter dan die van vergelijkbare leerlingen uit het huidige onderwijs?

Hier wordt er van uitgegaan dat een ander curriculum wordt gerealiseerd in de basisvorming. Dat zal dus allereerst moeten blijken. Daarna is pas de vraag zinvol of dit voor leerlingen maatschappelijk gesproken inderdaad een effect heeft. De criteria die daarbij worden aangelegd moeten voor een belangrijk gedeelte buiten het onderwijs worden gezocht. Men kan daarbij vooral denken aan arbeidsmarktpositie voor leerlingen na de basisvorming, de snelheid waarmee zij een baan krijgen, een geringere mate van werkloosheid of het verkrijgen van een werkkring op een hoger kwalificatieniveau dan vergelijkbare leerlingen zonder basisvorming. Het is duidelijk dat leerlingen niet alleen na hun vertrek uit de school moeten worden gevolgd, maar ook dat vele conjuncturele invloeden bij de feitelijke evaluatie meespelen.

D: *Vragen naar de vermindering van ongelijkheid van kansen door vermijding van te vroege studie- en beroepskeuze.*
D1: Vindt de keuze voor studie en beroep na afsluiting van de basisvorming, op een later moment plaats dan in het huidige voortgezet onderwijs?
D2: Neemt het boven onderzochte verschil in ongelijke kansen in studie- en beroepskeuze in de loop van de tijd meer af bij leerlingen met basisvorming dan bij soortgelijke leerlingen uit het huidige voortgezet onderwijs?

Bij de specificatie van deze evaluatievraag worden verschillende groepen leerlingen (sekse, milieu, prestatiegroepen en etnische achtergrond) onderscheiden. De veronderstelling is dat de prestaties in het onderwijs en in het beroep voor verschillende groepen niet groter, maar zelfs eerder kleiner zullen worden. Bovendien wordt verondersteld dat deze verschillen door een latere keuze nog verder zullen afnemen. Van belang is ook dat de specificaties betrekking hebben op de verschillende organisatievormen in het onderwijs, die een zelfstandig effect op een dergelijk criterium kunnen uitoefenen. Bovendien is van belang dat na eventuele 'kinderziekten' in de eerste jaren van de invoering van de basisvorming, in latere perioden gunstiger resultaten kunnen blijken.

E: *Vragen over de effecten op schoolorganisaties*
E1: Welke ervaringen worden opgedaan in de experimentele scholen bij de vormgeving van de basisvorming?
E2: Welke effecten heeft de invoering van de basisvorming op de structuur van het onderwijs zowel voor de organisatie als voor de leerlingen?

Deze vragen zijn vooral van *beschrijvende* aarde en worden van belang geacht voor de invoeringsstrategieën bij de basisvorming.

F: *Vragen die specifiek betrekking hebben op ontwikkelingen binnen de basisvorming*
F0: Welke resultaten worden er op de verschillende toetsen in de 14 vakken op de twee niveaus in de basisvorming gehaald?
F1: Zijn er belangrijke verschillen aan te wijzen in het tempo waarmee de leerlingen de toetsen basisvorming afleggen?
F2: Worden de verschillen in resultaten op de toetsen van de basisvorming tussen leerlingengroepen, scholen en organisatievormen in de loop van de tijd geringer?
F3: Welke vormgevingsaspecten spelen een rol binnen de basisvorming: varianten en organisatieproblemen in het realiseren van een nieuw curriculum?
F4: Hoe functioneren de eindtermen? Kunnen scholen er mee werken: welke effecten heeft dit voor curriculum en leerlingen?
F5: Welke verschillende lessentabellen worden gehanteerd en hoe wordt de vrije ruimte ingevuld: varianten en strategieën van scholen?

F6: Welke verschillende varianten en routes worden in het systeem van de basisvorming gehanteerd? In welke omvang komen de verschillende leerlingenstromen voor?

In de vragen onder F. wordt *geen* uitdrukkelijk vergelijkend gezichtspunt gehanteerd. Wel is het mogelijk om bijvoorbeeld d.m.v. tijdreeksen te laten zien dat bepaalde proporties in de loop van de tijd veranderen. Het is echter de vraag in hoeverre hier een evaluatief karakter aan moet worden toegekend. Niettemin is de informatie die op deze wijze wordt verzameld van belang voor de beoordeling van de invoering van de basisvorming. De gegevens hebben voor een belangrijk gedeelte een *beschrijvend* karakter.

5 *Het evaluatieplan*

5.1 *Voorstudies*

Het evaluatieplan voor de basisvorming sluit aan op een groot aantal beschikbare inzichten m.b.t. evaluatie in het onderwijs. Zo is er door Intomart (1986) een uitgebreid onderzoek onder schoolhoofden uitgevoerd, om na te gaan in hoeverre men bereid zou zijn aan een evaluatieplan mee te werken. Hieruit bleek dat 90% van de scholen bereid zou zijn te participeren: een duidelijk deel van de scholen echter gaf aan dat van deelname alleen sprake kon zijn als er niet van een overmatige belasting voor de scholen sprake zou zijn. Een aantal scholen stelde bovendien (soms geringe) aanvullende eisen m.b.t. de terugkoppeling van rapportages.

Door het CITO zijn verschillende voorstudies verricht op het terrein van de *instrumentenconstructie*. Zo werden itembanken aangelegd voor toetsen op de vier gekozen criteriumvakken. Reeds in een vroeg stadium was er een aparte studie uitgezet naar de beginkenmerken die over de schoolloopbaan moeten worden verzameld (Martinot, 1986).

Ook op het terrein van de *non-cognitieve* toetsen moet er nog een aanzienlijke hoeveelheid werk worden verricht. Allereerst is er geen sprake van een duidelijk gestandaardiseerde opvatting over wat er onder non-cognitieve kenmerken moet worden verstaan. Na een voorstudie door Doddema-Winsemius en Hofstee (1986) is een definitieve toets geconstrueerd die voor de basisvorming zal worden toegepast (Doddema-Winsemius, Haanstra & Pijl, 1987).

Een belangrijk 'gat' in de noodzakelijke kennis voor de evaluatie was het ontbreken van duidelijke clusters van variabelen op schoolniveau. In een apart vooronderzoek door het RION (Lugthart, 1988; Bosker, 1988) is nagegaan op welke wijze school-, inrichtings- en vormgevingskenmerken van het onderwijs en leerkrachten van belang kunnen zijn. De algemene conclusie van deze voorstudie is dat veel kenmerken tegenstrijdige indicaties opleveren. Niettemin is er een drietal clusters te identificeren – rondom het curriculum, stimulering en organisatie – die in het evaluatie-onderzoek kunnen worden gebruikt als controle op de vormgeving en institutionele structuur (een verdere detaillering geven Lugthart e.a. 1988, 186-189).

5.2 *De base-line meting*

Een belangrijk dilemma bij de evaluatie van de basisvorming was het volgende. Het hoofdaccent van de evaluatie zou liggen op een vergelijking tussen groepen in het huidige en het basisvormingsonderwijs. Tegelijkertijd zou echter worden nagegaan in hoeverre de ontwikkelingen in de loop van de tijd ten gunste van de basisvorming zouden gaan werken. Een belangrijk strategisch nadeel hierbij was dat er voor het huidige onderwijs op dit moment eigenlijk geen soortgelijke informatie beschikbaar is. Idealiter zou men gegevens willen hebben over een situatie die nog *niet gecontamineerd* is door de invoering van welk experiment dan ook. Zouden er in de loop van de tijd belangrijke veranderingen gaan plaatsvinden, dan zou men in ieder geval een soort 'nulmeting' beschikbaar moeten hebben. Een dergelijke meting vóórdat de evaluatie zelf van start zou gaan, zou bovendien een aantal andere voordelen kunnen hebben.

In de eerste plaats is er sprake van een relatief grootschalige evaluatie, dwars door alle onderwijsstructuren heen op elk niveau. Uitgangspunt van de evaluatie is dat er – op basis van goede instrumenten – vergelijkingen in de tijd moeten kunnen worden gemaakt. Daartoe dient eigenlijk een soort nul-meting beschikbaar te zijn. In het kader van de evaluatie is een dergelijke base-line door het CITO gezamelijk met het RION uitgevoerd, waaraan een viertal doeleinden is gekoppeld:

a. Op de vier gekozen terreinen van onderzoek is een breed samengestelde prestatie-

toets afgenomen aan een steekproef van ca. 30.000 leerlingen, verspreid over alle typen van het voortgezet onderwijs aan het eind van de derde klas. Op basis hiervan kunnen de definitieve toetsen worden geconstrueerd.

b. Nadat op basis van de breed samengestelde populatie de toetsen zijn geconstrueerd, worden de toetsscores voor deze populatie bepaald. Deze fungeren als een nul-meting voor de evaluatie van de basisvorming.
c. De niet-cognitieve toetsen zijn bij een deel van de steekproef afgenomen en fungeren hier als nul-meting.
d. Op schoolniveau is tegelijkertijd een aantal kenmerken van de schoolorganisatie en het curriculum verzameld, om tot een typologie te komen.

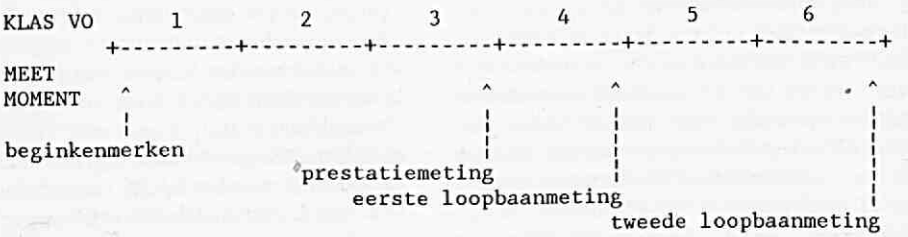
De base-line meting uit 1988 heeft aldus een tweetal verschillende doeleinden: in de eerste plaats worden de metingen verricht met het oog op de goede constructie van meetinstrumenten en de selectie van variabelen. In de tweede plaats worden de definitief gekozen en geconstrueerde kenmerken als een uitgangspunt voor verder onderzoek in de tijd vastgesteld.

5.3 *Het cohort-onderzoek*

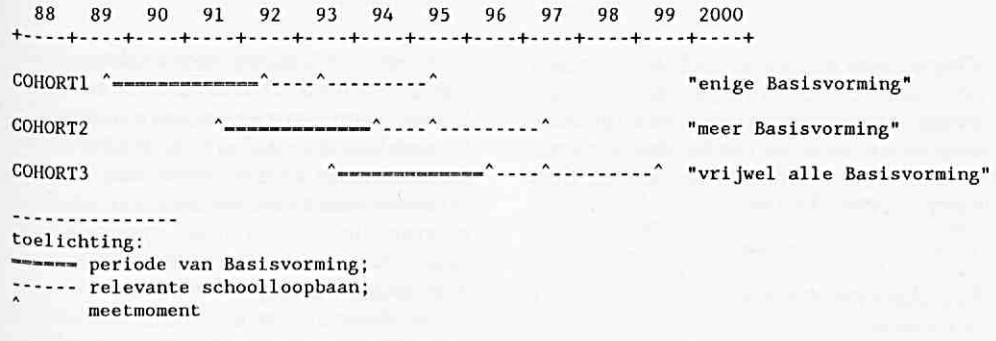
De verschillende vraagstellingen zullen voor een deel op basis van *verschillende* onderzoeken moeten worden beantwoord. Een gemeenschappelijk kenmerk in de meeste vraagstellingen is, dat informatie over leerlingen moet worden verzameld. De gegevensverzameling kan zodoende gemakkelijk op een centrale wijze plaatsvinden, waarbij nadien de informatie over evaluatievragen wordt uitgesplitst. De kern van het evaluatieprogramma is een samenstel van longitudinale cohorten van leerlingen. In Figuur 2 is de opzet van één moduul van het cohortonderzoek voor leerlingkenmerken weergegeven.

De eerste meting is aan het begin van het voortgezet onderwijs, terwijl de meting van afhankelijke variabelen feitelijk aan het eind van de derde klas zal plaatsvinden. Omdat sommige leerlingen langer over de basisvorming mogen doen (onder andere in de combinatievariant) zal echter minstens nog een tweetal andere metingen moeten plaatsvinden. Op deze momenten kan echter de gestandaardiseerde toets (zoals door het CITO ontwikkeld) niet meer worden afgenomen en moet volstaan worden met *loopbaanmetingen* in termen van behaald niveau van onderwijs of het gekozen beroep. Een dergelijk moduul is geschikt om de vraag A1 te beantwoorden. De evaluatie van A2-vragen is echter pas mogelijk als er meerdere modulen in de loop van de tijd worden onderzocht. Wanneer de tijd tussen de cohorten op twee jaar wordt gesteld, dan ontstaat er een drietal dakpansgewijze onderzoeken die informatie kunnen geven op de vraag in hoeverre verschillen in de loop van de tijd zijn verdwenen of juist toenemen.

Aanvullend is het mogelijk om ook twee eerder uitgevoerde CBS-cohorten (het zogenaamde SMVO- en het SLVO-cohort) aan de reeks toe te voegen zodat er uiteindelijk 5 cohorten over langere tijd kunnen worden vergeleken (periode 1977-1982-1989-1991-1993). Op het moment dat de onderzoekers op de scholen de leerlinginformatie verzamelen, wordt tegelijkertijd informatie over de scholen en curricula bijeengebracht. Zodoende kunnen in een *relationele database* de gegevens op leerling- en schoolniveau aan elkaar worden gekoppeld. De gegevens over de basisvormingstoetsen (voorzover relevant) moeten later via een bestandskoppeling met het CITO aan deze database worden toegevoegd.



Figuur 2 *Opzet van de module van het cohortonderzoek leerlingkenmerken*



Figuur 3 *Overzicht van de verschillende cohorten*

Een tweetal opmerkingen kan nog over deze gegevensverzamelingen worden gemaakt. In de eerste plaats is het waarschijnlijk dat de relevante set van variabelen en meetinstrumenten zo omvangrijk wordt dat deze niet in één grootschalig onderzoek kan worden opgenomen. Voor de evaluatie wordt daarom voorgesteld om uitsluitend de belangrijkste meetinstrumenten op te nemen. Aanvullende analyses en informatieverzamelingen kunnen later nog op geselecteerde scholen plaatsvinden (zie par 5.4). In de tweede plaats is er nog maar weinig bekend over de stabiliteit van allerlei schoolkenmerken. Eerder verricht onderzoek (zie Lugthart e.a. 1988) laat zien dat scholen die het ene jaar in termen van prestaties zeer effectief zijn, dat het andere jaar niet zijn. Om deze reden is het aan te raden schoolkenmerken *herhaald* te meten. Dit zou betekenen dat de 3 dakpansgewijze leerlingcohorten op dezelfde scholen moeten worden verzameld. Hierdoor kunnen de schooleffecten beter worden gemeten, terwijl tegelijkertijd een unieke mogelijkheid ontstaat om een studie te maken van de ontwikkelingen op scholen die naar de basisvorming overgaan. Op deze manier ontstaat dus een nieuwe dimensie van vergelijkend onderzoek (zie voor een uitwerking Bosker e.a. 1988).

De gekozen opzet voor de evaluatie impliceert een multi-level aanpak. Om deze reden is een voorstudie gestart om na te gaan op welke wijze de opslag van de gegevens het meest efficiënt zal kunnen gebeuren[3]. Vermeden moet immers worden dat er na de dataverzameling een lange periode voorbijgaat voordat de gegevens in een of andere vorm kunnen worden verwerkt. Daarnaast is een voorstudie gestart om na te gaan of de beschikbare multi-level analyseprocedures nog aanzienlijk kunnen worden verbeterd door meer efficiënte schattingsmethoden toe te passen, meerdere niveaus van analyse op te nemen en de mogelijkheid in te bouwen om gelijktijdig meervoudige criteria in de analyse op te nemen[4].

Organisatorisch levert een dergelijk stelsel van cohorten natuurlijk nog de nodige complicaties op. Zo is er nog geen duidelijkheid over de eigendomsrechten van de verschillende gegevens. Vanuit de coördinatie van de evaluatie is steeds benadrukt dat een onafhankelijke evaluatie moet inhouden dat ook na publikatie van evaluatieresultaten de data beschikbaar komen voor een *onafhankelijke contra-expertise* door vakgenoten. Dan kan binnen het wetenschappelijk forum effectief worden gecontroleerd of de juiste conclusies uit het onderzoek zijn getrokken[5].

5.4 *Dieptestudies*

De voorgaande benadering van de evaluatie is voornamelijk grootschalig en kwantitatief. Een belangrijk onderdeel van de evaluatie is echter dat ook op microniveau van scholen en leerlingen de relaties dieper moeten kunnen worden onderzocht. Er zijn verschillende mogelijkheden om dit gezichtspunt verder uit te werken. Enerzijds kan men een selectie van basisvormingsscholen nader ondervragen over de verschillende strategieën die zij hebben ondernomen om basisvorming in te voeren. Tegelijkertijd kunnen de resultaten van het grootschalig onderzoek hieraan worden gekoppeld om te zien of de strategieën ook inderdaad effectief zijn. Een andere mogelijkheid is om *uitbijters*, d.w.z. zowel de laag als hoog presterende leerlingen (in het reguliere en het basisvormingsonderwijs), te selecteren en apart in een diepte-onderzoek met kwalitatieve analyses op te nemen. Ten slotte is het ook mogelijk om door dieptestudies gedetailleerd na te gaan op welke punten basisvor-

mingsscholen en reguliere scholen in organisatorische kenmerken van elkaar verschillen, omdat de basisvorming nu eenmaal specifieke eisen m.b.t. kwaliteit van het docentencorps en de aantallen leerlingen vereist (vgl. Lugthart e.a. 1988, 188-189).

6 *Organisatorische context van de evaluatie*

Het is bekend dat organisatorische voorwaarden veel betekenen voor het welslagen van een onderzoeksprogramma. Dit geldt temeer voor evaluatie-onderzoek van een belangrijke onderwijsinnovatie, die het werkterrein van verschillende organisaties bestrijkt. In de eerste plaats heeft het Ministerie van Onderwijs met de beoogde wetgeving om een evaluatie gevraagd. Het Ministerie laat zich adviseren door de CCE. Deze bestaat uit vertegenwoordigers van het CITO, het CBS, de SVO en de onderwijsorganisaties, en is verder aangevuld met waarnemers van het Ministerie en de Onderwijsinspectie. De commissie adviseert aan het Ministerie over het wenselijke onderzoek en wordt voorgezeten door een onafhankelijke voorzitter (momenteel prof. dr. W. K. B. Hofstee). De vertegenwoordiging van het CITO ligt voor de hand; voor de evaluatie van de basisvorming moeten nieuwe toetsen worden geconstrueerd. De representatie van de SVO in de CCE is ingegeven door het feit dat alle onderwijsonderzoek in Nederland via de SVO zou moeten lopen. Het CBS is betrokken in de CCE vanwege haar ervaring met grootschalig cohortonderzoek. Voor de uitvoering van het evaluatieprogramma is door de minister jaarlijks een bedrag van 1 miljoen gulden in het vooruitzicht gesteld. Bovendien is een coördinator bij de SVO gedetacheerd, om de onafhankelijkheid van de evaluatie nogmaals te benadrukken.

Hoewel de vertegenwoordiging van alle relevante partners in de CCE een goede garantie lijkt voor de realisatie van het evaluatieprogramma, zijn tegelijkertijd enige potentiële spanningsbronnen ingebouwd. Zo heeft de SVO vanaf het begin af aan geprotesteerd tegen de instelling van de CCE. De opdrachtresearch op het gebied van het onderwijsonderzoek werd sinds de instelling van de Wet op de OnderwijsVerzorging immers exclusief toegedacht aan het Instituut voor Onderzoek van het Onderwijs: SVO. Hiermee is de advies- en beslissingsstructuur als het ware verdubbeld, hetgeen belangenconflicten in de hand werkt. Dit bleek in het verleden onder meer te spelen bij de afstemming met een andere grootschalige evaluatie – die van het onderwijsvoorrangsbeleid – waar een andere advies- en financieringsstructuur was gekozen.

De uitvoering van de verschillende cohortenonderzoeken loopt parallel aan de evaluatie van het OVB (zie Kloprogge, 1986 en de betreffende bijdrage in deze aflevering). Een groot aantal variabelen is immers algemeen en geldt voor beide evaluaties in het voortgezet onderwijs. Er kan dan ook een grotere efficiëntie en afstemming worden bereikt wanneer de kosten van de (omvangrijke) leerlingenonderzoeken kunnen worden gedeeld door twee evaluaties.

7 *Vooruitzichten*

De voorgenomen invoering van de basisvorming in het voortgezet onderwijs gedurende de komende 10 jaren, zou een ingrijpend proces van herstructurering zijn. Niet alleen zal er sprake zijn van een ander curriculum, waarbij voor grote groepen leerlingen meer leerstof in kortere tijd zal worden aangeboden, maar er zal ongetwijfeld ook sprake zijn van een groot aantal *onbedoelde effecten* van de invoering. Nu al is te voorzien dat er sprake zal zijn van vele fusies omdat de scholen op dit moment niet voldoen aan organisatorische en inhoudelijke eisen die bij de basisvorming worden gesteld (zie bijvoorbeeld Olthof, 1988). Door de open opzet van de evaluatie is het zeer goed mogelijk om na te gaan òf en op welke wijze schaalvergroting tot stand komt en welke problemen eventuele fusies van scholen tot gevolg hebben.

Een tweede onduidelijk punt bij de invoering van de basisvorming is dat op het moment dat de evaluatie moest worden gepland, nog niet bekend was òf en op welke wijze de basisvorming zou worden ingevoerd. De demissionaire status van het kabinet Lubbers II maakte de toekomst nog onzekerder: zouden de voorstellen worden gehandhaafd door het nieuwe kabinet? In het regeerakkoord van het kabinet Lubbers III is inmiddels vastgelegd dat op minstens één punt de voorgenomen

plannen worden gewijzigd: De basisvorming op *twee niveaus* behoort tot het verleden. Maar hoe de precieze uitwerking zal zijn staat nog evenmin vast. In ieder geval zitten we in de unieke situatie dat er een evaluatie is gestart vóórdat de onderwijshervorming wordt ingevoerd!

Het is een open vraag in hoeverre de invoering van de basisvorming in de komende jaren een mijlpaal zal zijn in de Nederlandse onderwijsgeschiedenis. De parlementaire behandeling moet nog plaats vinden en het lijkt nog allerminst zeker dat de – in meerdere opzichten – oorspronkelijke voorstellen van de WRR inderdaad zullen worden gerealiseerd. Niettemin is er in minstens twee opzichten sprake van een duidelijke winst. In de eerste plaats is er in vergelijking met de evaluaties uit de jaren *zeventig* sprake van een grote stap vooruit: er is de definitieve erkenning van het belang van evaluatie-onderzoek. In de tweede plaats kan het reeds begonnen evaluatie-onderzoek als een uitstekende start worden gezien van een zowel methodisch als inhoudelijk relevant onderzoekprogramma naar de effectiviteit van schoolkenmerken in het voortgezet onderwijs.

Het ware echter te wensen dat een dergelijke vooruitgang in de evaluatie van onderwijs gepaard zou gaan met een minstens even grote vooruitgang in het realiseren van *doelstellingen* in het onderwijs.

Noten

1. Dit is des te opvallender, omdat er regelmatig uitlatingen van de zijde van bewindslieden worden gehoord, die de wenselijkheid en noodzaak van evaluaties ontkennen. Zie bijvoorbeeld De Jager (1987), die op de opening van de ORD namens de minister sprak.
2. De nota is afzonderlijk gepubliceerd door SVO en is aldaar verkrijgbaar (Peschar, 1988).
3. Uitgevoerd door SCO, projectleider P. van Eeden.
4. Uitvoering door I. Kreft en J. de Leeuw, UCLA, Los Angeles.
5. Vergelijk de bekende discussie over het Coleman-onderzoek in 1966 en de Harvard seminars waarin de data opnieuw werden geanalyseerd. Zie daarvoor Mosteller en Moynihan (1968).

Literatuur

Beem, L.(red.), *Weten, Redeneren, Raden. Een analytische evaluatie van de Basisvorming*. Lisse: Swets & Zeitlinger, 1988.

Bosker, R. J., E. Lugthart & P. Roeders, *Schoolkenmerken en designs; verder algemene verkenningen*. Tweede deelrapport SVO project 6915. Groningen: RION, 1988.

Centraal Bureau voor de Statistiek, *Onderwijsindicatoren voor de basisvorming*. Voorburg: CBS, 1988.

Coördinatiecommissie evaluatieplan voortgezet onderwijs CCE, *Eerste advies, Uitvoeringsplan deel I*, Zoetermeer: Ministerie van O&W, 1985.

Coördinatiecommissie evaluatieplan voortgezet onderwijs CCE, *Vierde advies*, Zoetermeer: Ministerie van O&W, 1987.

Doddema-Winsemius, H. & W.K.B. Hofstee, *Operationalisering van het Onderwijsvoorrangsbeleid OVB*. 's-Gravenhage, SVO, 1985.

Doddema-Winsemius, H., F. Haanstra & Y. J. Pijl, *Exploratie niet-cognitieve toetsen ten behoeve van het evaluatieplan voortgezet onderwijs*. Groningen: RION, 1987.

Dronkers, J., De evaluatie van de maatschappelijke overwegingen om basisvorming in het onderwijs in te voeren. In: A. L. Beem (red.), *Weten, Redeneren, Raden. Een analytische evaluatie van de Basisvorming* (pp. 5-30). Lisse: Swets & Zeitlinger, 1988.

Intomart, *Haalbaarheid van de evaluatie van het voortgezet onderwijs*. Hilversum: Intomart, 1987.

Jager, A. de, Openingsrede van de minister van onderwijs en wetenschappen ter gelegenheid van de ORD 1987. In: W.K.B. Hofstee (red.), *Evaluatiemethodologie* (pp. 3-9). Lisse: Swets & Zeitlinger, 1987.

Kloprogge, J., *Evaluatie Onderwijsvoorrangsbeleid, Uitvoeringsplan*. 's-Gravenhage: SVO, 1986.

Kremers, E. J., *Vergelijking van leerresultaten in het voortgezet basisonderwijs en het vigerend onderwijs*. Arnhem: CITO, 1985.

Lugthart, E., P. J. B. Roeders, R. J. Bosker & K. T. Bos, *Effectieve schoolkenmerken in het voortgezet onderwijs: een literatuuroverzicht*. Groningen: RION, 1988.

Martinot, M., *Instrumenten beginmeting leerlingkenmerken: een voorbereidende studie in het kader van het evaluatieplan voortgezet onderwijs*. Arnhem: CITO, 1987.

Ministerie van Onderwijs en Wetenschappen, *Evaluatieplan voor het Voortgezet Onderwijs*. Zoetermeer: 1984.

Ministerie van Onderwijs en Wetenschappen, *Wetsontwerp Invoering Basisvorming*. 's-Gravenhage: SDU, 1987.

Ministerie van Onderwijs en Wetenschappen, Invoeringsplan Basisvorming. In: *Uitleg*, 2 maart 1988.

Mosteller. F. & D. P. Moynihan (Ed.), *On Equality of Opportunity*, New York: Vintage Books, 1968.

Olthof, A., Het invoeringsplan basisvorming, Doctoraalscriptie Vakgroep Sociologie, RU Groningen: 1988.

Peschar, J. L., *Evaluatie van de basisvorming. Kader voor het uitvoeringsplan*. 's-Gravenhage: SVO, 1988.

RION/ITS, *Voorbereiding evaluatie OVB voortgezet onderwijs* (SVO subsidieaanvraag). Groningen/Nijmegen: 1988.

Scheerens, J., *Evaluatie-onderzoek en beleid*. Harlingen: Flevodruk, 1983.

Scheerens, J., *Enhancing educational opportunities for disadvantaged learners*. Amsterdam: North-Holland Publishing Company, 1987.

Schouten, S. P., *Indicatoren ter evaluatie van de basisvorming op basis van CBS-gegevens*. Doctoraalscriptie Vakgroep Sociologie, RU Groningen: 1987.

Wetenschappelijke Raad voor het Regeringsbeleid WRR, *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij, 1986.

Curriculum vitae

J. L. Peschar is hoogleraar onderwijssociologie aan de Rijksuniversiteit Groningen. Van september 1986 tot oktober 1988 als evaluatie-coördinator betrokken bij de Evaluatie van de Eerste Fase Voortgezet Onderwijs (Basisvorming). Publiceerde onder meer over sociale ongelijkheid en onderwijskansen, met name in internationaal vergelijkend perspectief.

Adres: Sociologisch Instituut, Oude Boteringestraat 23, 9712 GC Groningen.

Manuscript aanvaard 6-12-'89

Summary

Peschar, J. L. 'With or without compulsary curriculum? Evaluation in secondary education in The Netherlands.' *Pedagogische Studiën*, 1990, *67*, 116-126.

In 1989 a large school reform in The Netherlands was scheduled: the introduction of a new compulsory curriculum (*basisvorming*) for the lower stage of secondary education. The evaluation of goals of the reform started already some years ago. Evaluation goals were translated into research questions and an extensive evaluation plan was designed in full agreement with the ministry of education. Although for long time it has been uncertain when the reform would become effective, the first stages of the evaluation have started yet. Three successive cohorts of students in both the traditional and the reform schooltype will be extensively tested and followed in their carreer. In addition qualitative and in depth studies into the curriculum and the school organisation are foreseen. The article discusses background and present stage of the evaluation and suggests that progress is made in two areas: the recognition of necessity of evaluation and the start of a large scale research programme on effective schools in secondary education.

De evaluatie van het onderwijsvoorrangsbeleid

M. P. C. VAN DER WERF
RION, Instituut voor Onderwijsonderzoek, Rijksuniversiteit Groningen

P. TESSER
ITS, Instituut voor Toegepaste Sociale Wetenschappen, Nijmegen

Samenvatting

In dit artikel wordt de opzet en uitvoering van de evaluatie van het Onderwijsvoorrangsbeleid (OVB) beschreven. Na een korte uiteenzetting over de doelen en de vormgeving van het OVB wordt de voorgeschiedenis van de evaluatie gememoreerd. Daarbij wordt ingegaan op de discussie die over het onderzoeksdesign, de effectmaten en de plaats van het implementatieonderzoek is gevoerd. Vervolgens komt de inhoud van het uitvoeringsplan voor de evaluatie aan de orde. Er wordt beschreven welke deelonderzoeken in dat plan zijn opgenomen en hoe de organisatie van de uitvoering is geregeld. Meer uitgebreid wordt aandacht besteed aan het hoofdonderzoek van de evaluatie, namelijk het onderzoek naar de implementatie en de effecten van het beleid. Besproken worden het steekproefdesign, het onderzoeksmodel, de gekozen variabelen en de uitvoeringsprocedure. Ten slotte worden enkele resultaten uit een vooronderzoek gepresenteerd en enige afsluitende conclusies getrokken.

1 *Inleiding*

In het begin van de jaren zeventig is naar aanleiding van publikaties over onderwijskansen van kinderen in achterstandssituaties (o.a. Van Heek, 1968) een begin gemaakt met de daadwerkelijke, systematische bestrijding van de onderwijsachterstanden van kinderen uit sociaal zwakke milieus. Aan scholen waarop veel van deze leerlingen zaten werden extra leerkrachten toegekend en enkele begeleidingsdiensten kregen extra subsidie. In de steden Amsterdam, Rotterdam en Utrecht werden projecten opgezet en programma's ontwikkeld voor onderwijs aan kansarme leerlingen.

In 1974 werd het 'Beleidsplan voor het onderwijs aan groepen in achterstandssituaties' uitgebracht, waarin een meer systematisch onderwijsstimuleringsbeleid werd voorgesteld dan tot dan toe was gevoerd. Op basis van dat beleidsplan kregen scholen extra faciliteiten, afhankelijk van hun achterstandsniveau, werden extra schoolbegeleiders en consulenten voor anderstaligen aangesteld en werden stimuleringsprojecten gestart, bestaande uit samenwerkingsverbanden van onderwijs- en welzijnsinstellingen. Langzamerhand werd de aandacht ook gericht op de bestrijding van onderwijsachterstanden van kinderen uit migrantengezinnen. In 1981 werd het beleidsplan 'Culturele minderheden in het onderwijs' uitgebracht. Toen reeds ontstond de wens om het onderwijsstimuleringsbeleid samen te voegen met het culturele minderhedenbeleid in wat men noemde het 'Onderwijsvoorrangsbeleid'. Deze wens werd in 1984 geconcretiseerd in een wetsvoorstel voor het Onderwijsvoorrangsbeleid (OVB) en kort daarna in het Onderwijsvoorrangsplan 1985-1989. Vooruitlopend op de parlementaire afhandeling van het wetsvoorstel is in september 1986 begonnen met het OVB. Door afwijzing van het wetsvoorstel in de Eerste Kamer is de wettelijke basis voor het beleid nog steeds niet gelegd. Wel is bepaald dat het beleid in ieder geval tot 1990 zal worden gevoerd overeenkomstig het wel goedgekeurde Onderwijsvoorrangsplan (zie voor een meer volledig historisch overzicht het ARBO-advies 'Voorrang aan achterstand', 1988).

In het Onderwijsvoorrangsplan is bepaald dat het beleid moet worden geëvalueerd om de effecten ervan vast te stellen. Vanaf 1984 zijn daarom diverse voorstudies verricht om te komen tot een acceptabel en uitvoerbaar evaluatieplan. Tijdens de voorbereiding van de evaluatie zijn discussies gevoerd over de te hanteren effectmaten, over de mate waarin aandacht moet worden besteed aan de implementatie van het beleid en over het evaluatie-

design. Voor een goed begrip van deze discussies is een korte uiteenzetting over de doelstellingen en vormgeving van het OVB noodzakelijk.

2 *Doelstellingen en vormgeving van het OVB*

De algemene doelstelling van het OVB wordt in het Onderwijsvoorrangsplan als volgt geformuleerd: 'Het Onderwijsvoorrangsbeleid is erop gericht door middel van een samenspel van gerichte maatregelen individuele scholen en welzijnsinstellingen, dan wel scholen en welzijnsinstellingen in een samenwerkingsverband verenigd, in staat te stellen onderwijsachterstanden van leerlingen van 4-18 jaar ten gevolge van sociale, economische en culturele omstandigheden op te heffen of te verminderen" (p. 5). Met het beleid worden in de eerste plaats cognitieve doelen nagestreefd, zoals het verwerven van kennis, inzicht en vaardigheden, wat dient te resulteren in een 'zo hoog mogelijk ontwikkelingsniveau'. Dit komt ook tot uitdrukking in het Ontwerp van wet op het Onderwijsvoorrangsbeleid en de Memorie van Toelichting daarop. Daarin staat het als volgt: 'De effecten van het OVB zullen in elk geval tot uitdrukking moeten komen in een verbetering van de ontwikkelingskansen en mede daardoor van de onderwijsleerprestaties van leerlingen in het primair en secundair onderwijs, waardoor mede de algemene doelstelling van dit onderwijs, met name voor leerlingen uit kansarme milieus, beter haalbaar wordt."

Het OVB bestaat uit twee componenten, te weten het gebiedenbeleid en het formatiebeleid. Het gebiedenbeleid houdt in dat gebieden met een cumulatie van sociale, economische en culturele problemen, faciliteiten krijgen voor coördinatie van voorrangsactiviteiten en voor het uitvoeren van specifieke projecten. Voorrangsgebieden bestaan uit samenwerkingsverbanden van 10 tot 20 scholen voor basis- en voortgezet onderwijs en welzijnsinstellingen. In totaal zijn 70 gebieden aangewezen als onderwijsvoorrangsgebied. Het belangrijkste criterium voor aanwijzing is de 'voorrangsscore' geweest. De voorrangsscore is de gemiddelde basisschoolscore van het gebied. De basisschoolscore wordt als volgt berekend: 0,75 × percentage leerlingen uit etnische minderheden + de SE-factor. De SE-factor geeft het percentage achterstandsleerlingen aan van de basisschool. Deze factor is opgebouwd uit enkele kenmerken van de leefsituatie van de leerling. De zo verkregen voorrangsscore kan per gebied variëren van minimaal 0 tot maximaal 275. De geselecteerde gebieden hebben alle een voorrangsscore van boven de 172. De gebieden hebben een gebiedsplan opgesteld dat moet functioneren als leidraad bij de uitvoering van de geplande activiteiten. De plannen beslaan gemiddeld een periode van vier jaar.

De formatiecomponent van het OVB houdt in dat aan individuele scholen waarop veel kinderen uit lagere sociale klassen en uit migrantengezinnen zitten extra personeelsformatie wordt toegekend. Hiertoe wordt aan de leerlingen een weegfactor gegeven: kinderen van ouders met een laag opleidings- en beroepsniveau krijgen een gewicht van 1.25 en kinderen uit migrantengezinnen een gewicht van 1.90. Ook leerlingen die in internaten of pleeggezinnen wonen en leerlingen van wie de ouders een trekkend bestaan leiden, krijgen een gewicht (respectievelijk 1.4 en 1.7). Leerlingen die niet tot één van de genoemde groepen behoren, krijgen een weegfactor van 1. Per school wordt het gewogen aantal leerlingen gedeeld door het werkelijke aantal en vermenigvuldigd met 100. Op basis van de resulterende score, die kan variëren van 100 tot 190, wordt de personeelsformatie (basisformatie genoemd) toegekend.

3 *Voorstellen voor het evaluatiedesign*

In het eerste voorstel voor de evaluatie van het OVB van Hofstee, Koster en Meijnen (1984) wordt het belang onderstreept van onderzoek naar de vraag of de met het beleid beoogde effecten, namelijk het opheffen of althans verminderen van de onderwijsachterstanden, inderdaad optreden. De evaluatie moet daarom in de eerste plaats gericht zijn op het peilen van onderwijsachterstand in de tijd. Hofstee et al. stellen dan ook voor op leerlingniveau herhaalde metingen van leerprestaties te verrichten bij opeenvolgende cohorten leerlingen in het tweede, vijfde en achtste leerjaar van het basisonderwijs en het derde leerjaar van het voortgezet onderwijs. De prestaties van de leerlingen uit achterstandsgroepen op OVB-

scholen dienen te worden afgezet tegen een normgroep van alle leerlingen, zowel door toetsing van verschillen tussen opeenvolgende leerjaren van dezelfde cohorten leerlingen als door toetsing van verschillen tussen opeenvolgende cohorten leerlingen in dezelfde leerjaren. Voor het samenstellen van de normgroep wordt een representatieve steekproef van 1 op 20 scholen voorgesteld, te trekken uit alle Nederlandse basisscholen en scholen voor voortgezet onderwijs. Deze steekproef wordt aangevuld met in beginsel alle scholen in de OVB-gebieden.

In hun reactie op het voorstel van Hofstee et al. maken Beem, Pijl en Scheerens (1986) enkele kritische kanttekeningen bij de voorgestelde longitudinale opzet waarbij herhaalde metingen op leerlingniveau zullen plaatsvinden. Aangezien Hofstee et al. niet alle cohorten in hetzelfde leerjaar laten starten, zijn de metingen over de factoren tijd en leerjaar slechts gedeeltelijk herhaald, hetgeen leidt tot problemen bij het ontwerpen van een statistisch model. Bovendien worden herhaalde metingen op leerlingniveau organisatorisch onaantrekkelijk en te kostbaar geacht. Als alternatief stellen Beem et al. daarom voor alleen herhaalde metingen op schoolniveau uit te voeren. Ook dan kunnen alle bovengenoemde verschillen worden getoetst. Wel dient dan gewerkt te worden met afzonderlijke statistische modellen voor basisscholen en scholen voor voortgezet onderwijs.

Een tweede kritische kanttekening die Beem et al. plaatsen bij het voorstel van Hofstee et al. heeft betrekking op het ontbreken van controlegroepen in het design, zodat causale uitspraken over beleidseffecten onmogelijk zijn. Zij pleiten daarom voor een quasi-experimenteel onderzoeksdesign waarin vier groepen scholen worden opgenomen, namelijk OVB gebiedsscholen, individueel gefaciliteerde scholen, niet-gefaciliteerde scholen en scholen die zijn afgewezen als gebiedsschool als (quasi)controlegroep. Deze laatste groep scholen moet voorrangsscores hebben die zoveel mogelijk gelijk zijn aan die van de geselecteerde groep.

In een reactie op het voorstel van Beem et al. stelt Hofstee dat het omvormen van de referentiegroep tot een (niet-equivalente, quasi) controlegroep leidt tot een ernstige verzwakking van het evaluatiedesign. Daarmee wordt immers de belangrijkste functie van de evaluatie voor de politieke en publieke discussie, namelijk het in de tijd peilen van onderwijsachterstanden, opgevat als discrepantie tussen OVB doelgroepleerlingen en landelijke referentiegroepen, opgeofferd (Hofstee, 1986).

De bezwaren tegen de voorstellen van Hofstee et al. enerzijds en die van Beem et al. anderzijds hebben geleid tot een evaluatieopzet waarin elementen uit beide voorstellen zijn gecombineerd tot het meest sterke en tegelijkertijd organisatorisch en financieel meest haalbare design (zie paragraaf 7.1).

4 *Discussie over de effectmaten*

In het eerste voorstel voor de evaluatie van het OVB stellen Hofstee et al. (1984) voor toetsen af te nemen voor het meten van onderwijsleerprestaties en intelligentie. De af te nemen toetsen voor onderwijsleerresultaten zouden minimale, communale leerdoelen moeten meten en zo 'onschools' mogelijk moeten zijn. Voorgesteld wordt in het basisonderwijs toetsen voor taal en rekenen en in het voortgezet onderwijs toetsen voor Nederlands, Engels en wiskunde af te nemen. De intelligentietests zouden een minimaal beroep moeten doen op aangeleerde, cultuurgebonden vaardigheden, en bij voorkeur geheel non-verbaal zijn. Als aanvullende effectmaat wordt voorgesteld doorstroom- en afstroomgegevens bij te houden.

In een latere studie werken Doddema-Winsemius en Hofstee (1985) de voorgestelde effectmaten verder uit. Met betrekking tot leerprestaties blijven de conclusies gelijk. De stelling dat er alleen non-verbale intelligentietests dienen te worden afgenomen, wordt genuanceerd. Aanvullend worden nu ook verbale tests genoemd. Hiermee komen zij tegemoet aan de wens van Jungbluth en Meijnen (1986) om de verbale tests zowel als zelfstandige effectvariabele als correctievariabele ten opzichte van leerprestaties te hanteren. Daarnaast wordt voorgesteld sociaal-emotionele effecten te meten. Als ankerbegrip hiervoor wordt 'sociale redzaamheid' gehanteerd, waarbij een onderscheid wordt gemaakt in vaardigheden en kennis. Het meten van houdingen wordt niet wenselijk geacht vanwege het probleem van indoctrinatie en vanwege de voorwendbaarheid van houdingen.

Het voorstel van Doddema-Winsemius en Hofstee vindt een brede acceptatie bij zowel onderzoekers, het onderwijsveld als de beleidmakers. Over de verwerping van 'houdingen' als effectmaat bestaat enige twijfel. Ten aanzien van dit punt wordt echter een pragmatisch standpunt ingenomen: als houdingen als tussendoelen worden nagestreefd of belangrijk lijken voor de verklaring van effecten wordt overwogen deze te meten.

De voorgestelde intelligentietests stuiten op grotere weerstand, vooral in het onderwijsveld. Uit een haalbaarheidsonderzoek van het RION en Intomart (Hoeben & Schaveling, 1986) blijkt dat veel schoolhoofden tegen afname van intelligentietests zijn. Deze weerstanden zijn waarschijnlijk te verklaren uit angst van scholen voor stigmatisering en uit een brede maatschappelijke doorwerking van de wetenschappelijke discussies over intelligentie-metingen (SVO, 1986). Ondanks deze weerstanden zijn de intelligentiemetingen in de voorgestelde vorm opgenomen in het evaluatieplan.

5 *De plaats van het implementatie-onderzoek*

In het programmavoorstel van Hofstee et al. (1984) wordt als doel van het implementatie-onderzoek genoemd het vaststellen van de wijze waarop de beleidsinstrumenten en -middelen op gebieds-, lokaal en schoolniveau worden ingezet. Alleen met behulp van implementatiegegevens is het mogelijk om gevonden effecten van het beleid te verklaren.

In aansluiting hierop hebben Jungbluth, Buis, Driessen en Mens (1985) een voorstel voor implementatie-onderzoek verder uitgewerkt. Zij komen tot verschillende deelonderzoeken: een breedteonderzoek naar de invoering van het OVB op schoolniveau, een diepteonderzoek op een beperkt aantal scholen, een beschrijvend onderzoek op gebiedsniveau en een exploratief diepteonderzoek bij scholen met extreme effectiviteitsscores.

Beem et al. (1986) maken bij deze voorstellen de volgende kanttekeningen: in de eerste plaats zou in het breedteonderzoek de 'mate van OVB-implementatie' gemeten moeten worden om daarmee de koppeling van het implementatie-onderzoek met de effectmeting te versterken. Daarbij zou gebruik moeten worden gemaakt van verschillende bestaande instrumenten voor schoolkenmerken. In de tweede plaats merken zij op dat diepteonderzoek van implementatievarianten bij scholen met extreme effectiviteitsscores o.a. kanskapitalisatie met zich meebrengt. Ten slotte stellen zij voor om door middel van een analytische evaluatie van het OVB hypothesen te genereren over de effectiviteit van bepaalde implementatievarianten.

In een nadere uitwerking van het eerste voorstel van Jungbluth et al. (1985) stellen Jungbluth en Meijnen (1986) dat de nadruk van het implementatie-onderzoek moet liggen op de breedtestudies. De breedtestudies op schoolniveau moeten het mogelijk maken om scholen vergelijkend te beschrijven op twee hoofdcategorieën van schoolkenmerken. Hiertoe onderscheiden zij de categorieën 'algemene schoolkenmerken' (waaronder kenmerken van effectief onderwijs) en 'typische onderwijsvoorrangsactiviteiten' (aanbodkenmerken voor achterstandsgroepen en in dit verband nagestreefde doelen). Op het meten van de 'mate van implementatie' en de koppeling van het implementatie-onderzoek aan het effectonderzoek wordt niet nader ingegaan. Betreffende het onderzoek bij scholen met extreme effectiviteitsscores merken Jungbluth en Meijnen op dat dit vooral is bedoeld om tekortkomingen in de evaluatieopzet op te sporen en om tot suggesties voor toekomstig beleid te komen.

Het uitvoeringsplan voor de evaluatie baseert zich grotendeels op de voorstellen voor implementatie-onderzoek van Jungbluth et al. (1985) en de uitwerking daarvan door Jungbluth en Meijnen (1986). De nadruk ligt op de breedtestudies op in eerste instantie schoolniveau en in tweede instantie gebiedsniveau. Voor de diepteстudies wordt verwezen naar de lange termijn. Over de koppeling van het implementatie-onderzoek op schoolniveau aan het effectonderzoek worden geen uitspraken gedaan. In paragraaf 7.1 wordt beschreven hoe deze koppeling feitelijk in de uitvoering van de evaluatie wordt gerealiseerd.

6 *Het uitvoeringsplan voor de evaluatie*

In het uitvoeringsplan voor de OVB-evaluatie (Kloprogge, 1986) worden als kernvragen genoemd:

a. De vraag of het beleidsprogramma leidt tot de beoogde effecten, door het opheffen of althans verminderen van onderwijsachterstanden ten gevolge van sociale, economische en culturele omstandigheden.
b. De vraag welke factoren in concreto verantwoordelijk zijn voor de mate waarin het programma effect sorteert.

Het centrale deel van de evaluatie bestaat uit longitudinale onderzoeken in het basis- en voortgezet onderwijs. In deze onderzoeken worden tweejaarlijks testen voor het meten van leerprestaties, intelligentie en sociale redzaamheid afgenomen en worden gegevens verzameld over de implementatie van het OVB. De onderzoeken in het basisonderwijs zijn gestart in september 1988, nadat door het ITS en het RION de voorbereidingen daartoe waren getroffen. In de volgende paragraaf wordt meer informatie verstrekt over de opzet en de wijze van uitvoering. Op dit moment kunnen nog geen resultaten worden gerapporteerd. De onderzoeken in het voortgezet onderwijs zijn gestart in september 1989. De uitvoering van dit onderzoek geschiedt in samenhang met de uitvoering van de evaluatie van de basisvorming in het voortgezet onderwijs.

Ten behoeve van de longitudinale onderzoeken in het basisonderwijs is in 1987 begonnen met de constructie van toetsen. Door het CITO zijn reken- en taaltoetsen ontwikkeld (Van Bergen, 1989). De testen voor het meten van intelligentie en sociale redzaamheid zijn ontwikkeld door het RION (Doddema-Winsemius & Van der Werf, 1988).

In 1987/1988 is een oriënterend onderzoek uitgevoerd in het basis- en voortgezet onderwijs ter vaststelling van de beginsituatie van de scholen en gebieden en van de wijze waarop op de scholen en in de gebieden het OVB is geïntroduceerd. In hetzelfde onderzoek zijn ook de instrumenten voor het meten van de implementatie van het OVB uitgeprobeerd. Het onderzoek is uitgevoerd door het ITS en het RION gezamenlijk. Verderop worden enkele resultaten gerapporteerd (zie paragraaf 8).

Eveneens in 1987/1988 is een analyse gemaakt van de gebiedsplannen van de onderwijsvoorrangsgebieden met als doel na te gaan of de gebiedsplannen goede plandocumenten zijn, of de doelen en voorgenomen activiteiten van de gebieden overeenkomen met de beleidsdoelen van het OVB en wat de verschillen en overeenkomsten zijn tussen de gebieden. Dit onderzoek is eveneens door het RION en het ITS gezamenlijk uitgevoerd. Voor de resultaten van dit onderzoek wordt verwezen naar Vermeulen en Van der Werf (1989).

Naast deze onderzoeken is een aantal specifieke projecten voorzien of reeds in uitvoering: een onderzoek in enkele gebieden, een onderzoek naar de relatie tussen OVB en de basiseducatie, een onderzoek naar de percepties en aspiraties van buitenlandse gezinnen, een onderzoek naar specifieke projecten voor buitenlandse meisjes, een onderzoek naar het OVB in het speciaal onderwijs, een onderzoek naar de ondersteuning, een onderzoek naar ervaringen en problemen van de betrokkenen bij OVB.

Voor de evaluatie is 1.5 miljoen gulden per jaar uitgetrokken door het Ministerie van O & W. De evaluatie wordt gecoördineerd door het Instituut voor Onderzoek van het Onderwijs (SVO). Het centrale deel van de evaluatie wordt gezamenlijk uitgevoerd door het ITS en het RION. Specifieke onderzoeken worden aanbesteed bij andere onderzoeksinstituten of universiteiten. Een projectgroep, ingesteld door de Minister van O & W, treedt op als intermediair tussen de onderzoekers en de beleidmakers. Alle onderzoeksvoorstellen worden beoordeeld door een externe SVO beoordelingscommissie.

7 *Het longitudinale implementatie- en effectonderzoek in het basisonderwijs*

7.1 *Onderzoeksontwerp*

In de longitudinale onderzoeken in het basisonderwijs worden zowel de implementatie als de effecten van het OVB onderzocht. De onderzoeksopzet kan worden gekarakteriseerd als een combinatie van quasi-experimenteel onderzoek en longitudinaal cohortonderzoek. Vanwege de organisatorische en financiële haalbaarheid van de evaluatie worden echter geen herhaalde metingen op leerlingniveau uitgevoerd, maar worden scholen of liever gezegd klassen gevolgd. Het gaat hier om de klassen 4, 6 en 8. Deze klassen worden tweejaarlijks getest. Binnen de klassen kan telkens een groot deel van de leerlingen worden teruggevonden, maar ten gevolge van tussentijdse in- en uitstroom van leerlingen kan

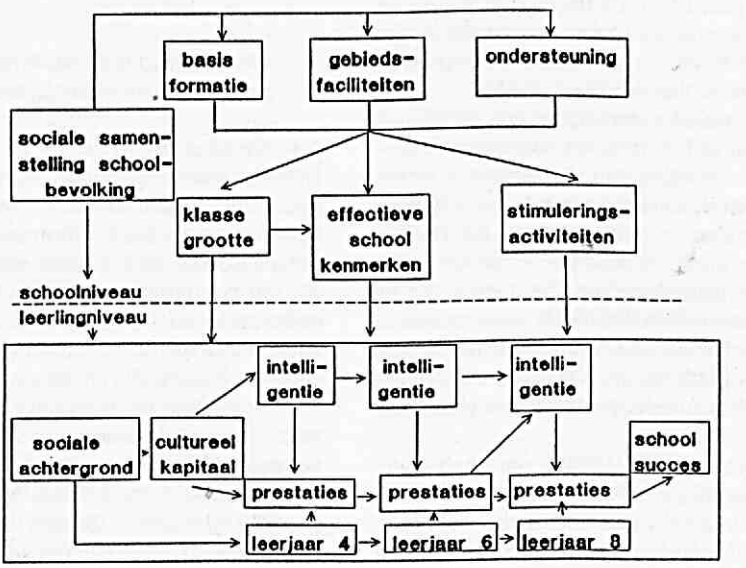
de klassesamenstelling veranderen.

Het steekproefdesign ziet er als volgt uit: de kern van de steekproef is een 3,5% random steekproef uit alle basisscholen. Dit is de door Hofstee et al. (1984) gewenste referentiegroep. Deze kernsteekproef is gestratificeerd op basis van de gebiedenfactor en de schoolscorefactor. De gebiedenfactor heeft twee niveaus, scholen die tot een gebied behoren en niet-gebiedsscholen. De schoolscores weerspiegelen de sociale en etnische samenstelling van de scholen en daarmee de hoeveelheid formatie op basis van de formatiecomponent van het OVB. Ze hebben een bereik van 100 tot 190 (zie paragraaf 1). Relevante scheidingspunten zijn 105 en 115. De scheidingspunten resulteerden in een factor met drie niveaus. Deze stratificatie van de kernsteekproef leverde zes strata op. Strata met minder dan 100 scholen in de kernsteekproef zijn aangevuld tot honderd. De resulterende steekproef bestaat uit 826 scholen (zie voor een meer uitgebreide beschrijving van de wijze van steekproeftrekking Tesser & Mulder, 1987).

Dit design maakt het enerzijds mogelijk uitspraken te doen over de effecten van het OVB in de tijd door de prestaties van de leerlingen uit de doelgroepen van het OVB af te zetten tegen die van de referentiegroep. Anderzijds biedt het design de mogelijkheid verschillen te toetsen tussen gebiedsscholen en niet-gebiedsscholen en tussen scholen met geen of nauwelijks, redelijk veel en veel extra formatie op basis van de OVB formatiecomponent.

Het onderzoek naar de implementatie van het OVB wordt uitgevoerd op dezelfde scholen als waar het effectonderzoek plaatsvindt. Daartoe worden op zowel school- als klasseniveau gegevens verzameld. Op deze wijze kunnen de gegevens over de implementatie van het OVB worden gekoppeld aan het effectonderzoek. Voor het opsporen van factoren waaraan de eventuele effecten van het OVB kunnen worden toegeschreven is een multilevel model ontwikkeld waarin de relaties tussen variabelen op schoolniveau en variabelen op leerlingniveau worden beschreven. Dit model ziet er als volgt uit:

Het model heeft twee niveaus, een schoolniveau en een leerlingniveau. In het model wordt verondersteld dat schoolvariabelen effect hebben op de parameters van het individuele model. Schooleffecten zijn dan effecten op de intercepten en slopes van de regressies van prestaties op gezinsfactoren, sociale achtergrond etc. binnen scholen. Op basis van de



Figuur 1 *Multi-level model voor verschillen in schoolsucces*

literatuur over schooleffectiviteit en onderwijs voor kinderen uit achterstandssituaties wordt verondersteld dat leerlingprestaties beïnvloed kunnen worden door school- en instructiekenmerken die positief samenhangen met de leerprestaties van alle leerlingen en door specifieke stimuleringsactiviteiten voor achterstandsleerlingen. Materiële en personele faciliteiten en extra begeleiding hebben alleen effect via deze variabelen. Klassegrootte zal hetzij een direct, het zij een indirect – via de kenmerken van effectieve instructie – effect hebben. De onderzoeksresultaten tot nu toe geven hierover geen uitsluitsel. De situatie wordt verder gecompliceerd door de situatie dat scholen deel kunnen uitmaken van een onderwijsvoorrangsgebied (op basis van de gebiedencomponent van het OVB). In feite brengt de gebiedencomponent een nieuw niveau in het model. Voorlopig beperken we ons echter door de gebiedencomponent te behandelen als een extra OVB facilitering die via de kenmerken van effectief onderwijs en de stimuleringsactiviteiten effect zal hebben (zie voor een meer uitgebreide bespreking van het model en van de variabelen Tesser & Van der Werf, 1989).

7.2 *Uitvoeringsprocedure*

In mei 1987 zijn de scholen uit de steekproef schriftelijk benaderd met informatie over de evaluatie en het verzoek medewerking daaraan te verlenen. Behalve de schooldirecteuren is ook aan de besturen van de scholen informatie verstrekt. Daarnaast is contact geweest met de coördinatoren van de betrokken gebieden. In juni/juli zijn de scholen gebeld. Aan de scholen die aan de evaluatie wilden meewerken is gevraagd een aantal gegevens te verstrekken over aantallen leerlingen, combinatieklassen, parallelklassen e.d. Indien een school weigerde werd deze vervangen door een school uit het overeenkomstige scholenstratum.

In augustus zijn de scholen die bereid waren mee te doen aan de evaluatie gebeld door de testleiders die deze scholen gingen bezoeken voor het maken van afspraken. De testleiders hadden van tevoren een uitgebreide instructie en een training voor de testafname gehad. Alle testleiders hadden enige jaren onderwijservaring en/of reeds ervaring met het klassikaal afnemen van toetsen. Elke school is in principe drie dagen bezocht voor het afnemen van de testen voor rekenen, taal, intelligentie en sociale redzaamheid in groep 4, 6 en 8 en voor het inzamelen van de al eerder opgestuurde vragenlijsten voor directeuren en leerkrachten. De taaltesten werden voorafgegaan door een vragenlijstje voor het verkrijgen van achtergrondgegevens over de leerlingen. Aan de leerlingen werden vragenlijsten aan de ouders meegegeven voor het verzamelen van gegevens over opleiding en beroep, deelname aan culturele activiteiten en onderwijsstimulerend gedrag. Deze vragenlijsten zijn via de school opgestuurd aan de onderzoekers. Het was de bedoeling de dataverzameling voor de kerstvakantie af te ronden. Dat bleek in de praktijk niet haalbaar. Begin maart 1989 zijn de laatste gegevens binnengekomen.

Aan de scholen is een rapportage over de toetsresultaten toegezegd. Daarin worden de toetsresultaten voor taal en rekenen van elke leerling afgezet tegen het landelijk gemiddelde. Na afhandeling van de schoolrapporten kon in september 1989 worden begonnen met de analyse van de gegevens ten behoeve van de rapportage over de eerste fase van de longitudinale onderzoeken in het basisonderwijs.

8 *Enkele resultaten uit het oriënterend onderzoek*

8.1 *Opzet en uitvoering van het oriënterend onderzoek*

Het oriënterend onderzoek in het basisonderwijs is opgezet overeenkomstig de longitudinale onderzoeken. Er zijn gegevens verzameld over leerlingen en over school- en klassekenmerken. De gegevens over school- en klassekenmerken zijn verkregen met behulp van deels schriftelijke, deels mondelinge vragenlijsten bij directeuren en bij de leerkrachten van groep 4, 6 en 8. De school- en klassegegevens hebben betrekking op materieelstructurele schoolkenmerken (b.v. schoolgrootte, hoeveelheid formatie, e.d.), kenmerken van effectief onderwijs op school- en klasseniveau, stimuleringsactiviteiten en de wijze van introductie van het OVB op de scholen. De gegevens over leerlingen gaan over het leerlingcohort 1986, dat wil zeggen de leerlingen die in het schooljaar 1986-1987 in het laatste leerjaar van de onderzochte basisscholen hebben gezeten. De belangrijkste gegevens be-

treffen het advies voor de keuze van het schooltype in het voortgezet onderwijs, het leerlinggewicht voor de formatieregeling en het beroep en de opleiding van de ouders. De leerlinggegevens zijn overgenomen uit de leerlingadministraties van de onderzochte scholen.

Er is gewerkt met een steekproef van basisscholen conform de indeling in groepen scholen zoals in het longitudinale onderzoek (zie par. 7.1). De dataverzameling heeft plaatsgevonden in de periode februari-april 1988. De scholen zijn eerst per brief op de hoogte gesteld van het onderzoek en vervolgens telefonisch benaderd voor medewerking. Van de oorspronkelijk aangeschreven scholensteekproef heeft uiteindelijk ruim 50 procent aan het onderzoek meegedaan. De uitgevallen scholen zijn vervangen door reservescholen uit dezelfde groep. In totaal zijn er gegevens verkregen over 202 scholen en 3623 leerlingen. Weigering tot deelname aan het oriënterend onderzoek heeft zich bij scholen in voorrangsgebieden aanzienlijk minder voorgedaan dan bij niet-gebiedsscholen. Van de gebiedsscholen uit de oorspronkelijke steekproef heeft 65 procent meegedaan. Van de niet-gebiedsscholen 48 procent. Vergelijking van de adviezen voor voortgezet onderwijs van de leerlingen van de scholen uit de referentiesteekproef met landelijke gegevens wijst uit dat de opgetreden uitval de representativiteit van de referentiesteekproef niet ernstig heeft aangetast.

In de volgende paragrafen wordt gerapporteerd over de onderwijsachterstanden van de OVB-doelgroepen, over verschillen hierin tussen scholen en over de perceptie en de betekenis van onderwijsvoorrang op schoolniveau.

8.2 *Achterstanden van de OVB doelgroepen en verschillen tussen scholen*

Ten behoeve van de evaluatie van het OVB is het interessant om na te gaan hoe het na een jaar uitvoering van het beleid staat met de onderwijsachterstanden van de doelgroepen van het OVB. Voor de beantwoording van deze vraag beschikken we over de adviezen voor voortgezet onderwijs van het cohort leerlingen dat in het schooljaar 1986/1987 in het laatste leerjaar van het basisonderwijs zat. Deze leerlingen hebben dus een jaar het OVB meegemaakt. Op basis van de weegfactoren uit de OVB-formatieregeling is de cohort 86 steekproef verdeeld in drie hoofdgroepen: de 1.90-leerlingen, de 1.25-leerlingen en de 1.00-leerlingen. Om uitsluitsel te krijgen over de onderwijsachterstanden van de OVB doelgroepen is gekeken naar de relatie tussen weegfactor en advies voor de keuze van het schooltype in het v.o. De gegevens staan in Tabel 1.

De kans op een LBO-advies is bij de OVB-doelgroep leerlingen 2,5 keer zo groot als bij de overige leerlingen. De omgekeerde trend geldt voor de HAVO-VWO-adviezen. Vergelijken we de OVB doelgroep leerlingen met de referentiegroep, dan blijkt de kans op een LBO-advies voor de 1.25 en de 1.90 leerlingen 1,5 keer zo groot als voor de 'gemiddelde leerling' in het basisonderwijs. De kans op een VWO-advies is voor de 1.25 leerlingen de helft en voor de 1.90 leerlingen meer dan tweederde kleiner dan voor de leerlingen uit de referentiegroep.

Opmerkelijk zijn de geringe verschillen tussen de 1.25 en de 1.90 leerlingen. Hoewel de gegevens over herkomstland, beroepsniveau en opleidingsniveau van de ouders door het grote aantal ontbrekende gegevens

Tabel 1 *Adviezen van de leerlingen uit cohort 86 naar gewichtscategorie (in procenten)*

	LBO	MAVO	HAVO	VWO
1.00-leerlingen*	17	29	28	26
1.25-leerlingen**	43	34	15	9
1.90-leerlingen**	43	33	19	5
alle leerlingen*	28	33	22	17

* referentiesteekproef
** totale steekproef

verre van compleet zijn, bestaan er aanwijzingen dat wanneer rekening wordt gehouden met deze drie factoren de verschillen tussen beroepsniveau en tussen herkomstland wegvallen. De verschillen in adviezen tussen de 1.25 leerlingen en de 1.90 leerlingen kunnen dan volledig worden toegeschreven aan de verschillen in opleidingsniveau van de vaders van de leerlingen (zie Tesser, Van der Werf, Mulder & Weide, 1989).

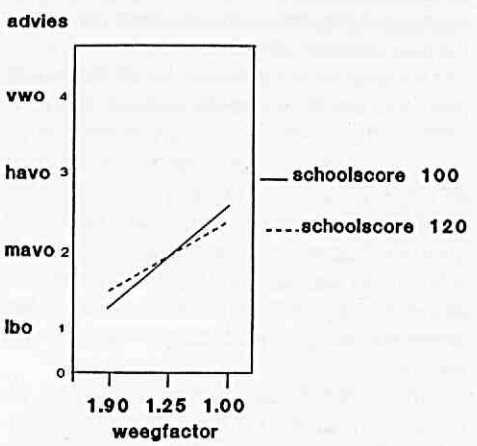
Uit deze gegevens blijkt dat het OVB in het algemeen na een jaar nog niet veel effect heeft gehad op de adviezen van de doelgroep leerlingen voor het voortgezet onderwijs. Het is echter mogelijk dat ten aanzien van deze adviezen wel verschillen tussen scholen bestaan.

De relaties tussen verschillen op school- en leerlingniveau in het cohort 86 onderzoek is uitgezocht met behulp van variantiecomponentenanalyse (Longford, 1987). In dit type analyse wordt het advies van een leerling opgesplitst in drie componenten: het gemiddelde advies van alle leerlingen uit cohort 86, het gemiddelde van de school waar de leerling op zit als afwijking van het landelijk gemiddelde en de mate waarin het advies van de leerling afwijkt van het schoolgemiddelde. Van de laatste twee componenten worden de varianties geschat. De verhouding tussen de varianties geeft aan in welke mate verschillen in advies kunnen worden toegeschreven aan scholen dan wel aan leerlingen.

De schatting van de variantie van de tussen scholencomponent van de adviezen van cohort 86 komt uit op .049, die van de tussen leerlingencomponent op .996. De verschillen in advies tussen leerlingen binnen scholen zijn vele malen groter dan de verschillen tussen scholen. Nog geen vijf procent van de totale variantie in advies is tussen-scholenvariantie. De conclusie dat onderwijsresultaten van leerlingen veel meer door leerlinggebonden buitenschoolse factoren worden bepaald dan door binnenschoolse factoren ligt voor de hand.

Toch moeten we hiermee uitkijken. De tussen-scholenvariantie heeft betrekking op de gemiddelden van scholen en is dus 'gespreid' over alle leerlingen. Kleine veranderingen in de tussen-scholencomponent kunnen daardoor binnen de totale groep leerlingen een relatief groot effect hebben. Voor het OVB is vooral interessant in welke mate verschillen in advies tussen scholen worden bepaald door de sociale achtergrond van de leerlingen. Om daar iets over te kunnen zeggen is in het analysemodel een effect op individueel niveau van de weegfactor opgenomen. De tussen scholen- en tussen-leerlingenvarianties veranderden daardoor aanzienlijk. De tussen-scholenvariantie daalt van .049 naar .032, een afname met ongeveer 35 procent. Zo'n 35 procent van de verschillen tussen scholen kan worden toegeschreven aan een op alle scholen gelijk effect van de weegfactor.

Dit betekent dat het effect van de weegfactor niet op alle scholen gelijk is en dat sommige scholen dus beter in staat zijn te compenseren voor het sociaal en cultureel milieu van de leerlingen dan andere scholen. De vraag die dan rijst is of dit een effect is van het OVB. Met andere woorden: leidt de extra personele ruimte die scholen krijgen in het kader van het OVB tot hogere adviezen voor voortgezet onderwijs voor de doelgroepleerlingen? Scholen met een hoge schoolscore hebben meer personele ruimte dan scholen met een lage schoolscore. De extra personele ruimte wordt gedeeltelijk gebruikt voor het verkleinen van de klassen, gedeeltelijk voor andere activiteiten. Wanneer de schoolscore, het aantal leerlingen per formatieplaats en de gemiddelde klassegrootte worden ingebracht in het analysemodel, dan blijkt dat de schoolscore wel effect heeft op de adviezen, maar de klassegrootte en het aantal leerlingen per formatieplaats niet. In Figuur 2 zijn de relaties tussen de ad-



Figuur 2 *Relatie tussen weegfactor en advies bij verschillende schoolscores*

viesscores en de weegfactoren op twee scholen afgebeeld. De ene school heeft een schoolscore van 100, de andere een schoolscore van 120. Het gaat hierbij om schattingen op basis van de analyseresultaten. Uit Figuur 2 valt af te lezen dat de 1.90 leerlingen hogere adviezen krijgen naarmate de schoolscore hoger is. Voor de 1.00 leerlingen geldt het omgekeerde.

Uit nadere analyse (zie Tesser et al., 1989) blijkt echter dat deze effecten niet eenduidig geïnterpreteerd kunnen worden. De oorzaak van de gevonden effecten is vermoedelijk de wat hogere sociale achtergrond van de 1.00 leerlingen op scholen met een lage schoolscore. Het ontbreken van significante effecten van de personele ruimte en de groepsgrootte is opmerkelijk. Hieruit blijkt dat in de beginfase van het OVB het beschikken over extra personele ruimte en de daarmee samenhangende kleinere klassen geen effect heeft op de adviezen van de leerlingen.

In de volgende paragraaf zal duidelijk worden dat dit vermoedelijk voor een belangrijk deel kan worden toegeschreven aan de geringe mate waarin het OVB in de scholen leeft.

8.3 *Het voorrangsbeleid in de scholen*

In het oriënterend onderzoek zijn op schoolniveau gegevens verzameld over de kennis en perceptie van het OVB en over de wijze waarop in de scholen aan het OVB vorm wordt gegeven. Uit de resultaten komen grote verschillen naar voren tussen de scholen in voorrangsgebieden en niet-gebiedsscholen. Daarbij moet worden bedacht dat het merendeel van de niet-gebiedsscholen in het onderzoek percentages doelgroepleerlingen heeft die vergelijkbaar zijn met de gebiedsscholen.

Directeuren van gebiedsscholen weten veel meer van het OVB dan directeuren van niet-gebiedsscholen. Ruim veertig procent van de directeuren van niet-gebiedsscholen zegt niet of nauwelijks op de hoogte te zijn van het OVB. Dit percentage wordt niet beïnvloed door de hoeveelheid extra formatie die de scholen als gevolg van het OVB-formatiebeleid krijgen. Bij de gebiedsscholen is het overeenkomstige percentage slechts zes procent.

Hoe slecht de niet-gebiedsscholen op de hoogte zijn van het OVB wordt vooral duidelijk uit de beantwoording van de vraag of de school extra formatie krijgt in het kader van OVB. Van de niet-gebiedsscholen die feitelijk extra OVB-formatie krijgen, blijkt niet meer dan dertien procent dit ook te weten. Van de gebiedsscholen is er daarentegen vrijwel geen enkele directeur die niet weet dat zijn school extra OVB-formatie krijgt. Wanneer directeuren niet eens weten dat ze extra formatie krijgen, wordt het wel zeer onwaarschijnlijk dat die formatie wordt ingezet voor het doel waarvoor ze bestemd is.

Aan de directeuren die wel weten dat ze extra formatie krijgen ten gevolge van het voorrangsbeleid, is gevraagd wat er met die formatie wordt gedaan. Ook op dit punt zijn er grote verschillen tussen scholen die deel uitmaken van voorrangsgebieden en niet-gebiedsscholen. Op gebiedsscholen wordt de extra formatie veel meer gebruikt voor specifieke activiteiten, gericht op verbetering van de onderwijskansen van de OVB-doelgroepleerlingen. Op niet-gebiedsscholen wordt de extra formatie veel meer gebruikt voor groepsverkleining.

Ook in hun betrokkenheid bij OVB-aangelegenheden verschillen de scholen binnen en buiten de gebieden. Publikaties over het OVB, zoals het blad Stimulans, de OVB-brochures en artikelen over het OVB, worden meer gelezen op gebiedsscholen dan op de overige scholen. Verder nemen deze scholen vaker deel aan onderwijsvoorrangsactiviteiten, zoals OVB-themadagen, OVB-werkgroepen en OVB-activiteiten van de onderwijsbegeleidingsdienst. Aan OVB-nascholing wordt op bijna geen enkele school deelgenomen.

Wat betreft de houding tegenover het OVB zijn er niet veel verschillen tussen de scholen. De scholen die op de hoogte zijn van het beleid zijn er vrij positief over. Met uitspraken als "De doelstellingen van het OVB spreken mij aan", "Het OVB is een goede zaak" en "Het OVB moet worden voortgezet" is vrijwel iedereen het eens. Wel hebben gebiedsscholen wat meer vertrouwen in de mogelijkheden en effecten van het OVB dan de overige scholen. Ook over de faciliteitenregeling van het OVB hebben de scholen die de regeling kennen in het algemeen een positief oordeel. Vooral gebiedsscholen vinden de regeling noodzakelijk, waardevol en zinvol. Bijna alle scholen zijn het er echter over eens dat de faciliteitenregeling (te) beperkt is.

9 *Conclusies*

De voorbereiding van de evaluatie OVB heeft een vrij lange periode in beslag genomen. Nu al kan worden gezegd dat die lange voorbereiding een gunstig effect heeft gehad op de uitvoering van de evaluatie. Dankzij de discussies die gevoerd zijn over de opzet was het mogelijk aan de evaluatie te beginnen met een goed doordacht en uitgewerkt plan. Dit heeft de steekproeftrekking en het uitwerken van de uitvoeringsprocedure aanzienlijk vergemakkelijkt. Dankzij de discussies over de effectmaten en de uitwerking van de operationalisering daarvan konden de af te nemen tests tijdig worden opgeleverd. Misschien ook dankzij deze voorbereiding waren de weerstanden, die begin 1986 nog tegen het afnemen van tests, met name van intelligentietests leken te bestaan, drastisch afgenomen. In de vooronderzoeken die in 1987 zijn uitgevoerd, met name het oriënterende onderzoek, zijn vervolgens de instrumenten ontwikkeld en beproefd voor het verzamelen van gegevens over school- en klassekenmerken en over specifieke stimuleringsactiviteiten. Ook is in het oriënterend onderzoek ervaring opgedaan met procedures om scholen te benaderen en gegevens over leerlingen te verzamelen. Ten aanzien van dit laatste is gebleken dat scholen zeer terughoudend zijn in het verstrekken van gegevens over opleiding en beroep van ouders. Mede daarom is besloten om deze gegevens via de ouders zelf te verzamelen. Al deze voorbereidingen tezamen hebben het mogelijk gemaakt dat de eerste ronde van dataverzameling in het basisonderwijs redelijk succesvol is verlopen.

Uit de resultaten van het oriënterend onderzoek die hier zijn gerapporteerd blijkt duidelijk dat het OVB nog lang niet op alle scholen die faciliteiten ontvangen is doorgedrongen. Veel scholen weten niet eens dat de extra basisformatie die ze ontvangen op grond van de weegfactoren bedoeld is voor het verminderen van de onderwijsachterstanden van de doelgroepen van het OVB. Het hoeft dan ook geen verbazing te wekken dat de extra personele ruimte die scholen krijgen ten gevolge van deze regeling geen effect heeft op de leerlingen. Uit deze constatering blijkt eens te meer het nut van het implementatieonderzoek en de koppeling van dit onderzoek aan het effectonderzoek.

Ten slotte blijkt uit het oriënterend onderzoek dat effecten van het OVB pas op langere termijn dan een à twee jaar zijn te verwachten. Het longitudinale evaluatiedesign, waarin tweejaarlijks de onderwijsachterstanden van de doelgroepen van het OVB worden gepeild, lijkt daarom voorshands de meest geëigende onderzoeksopzet te zijn voor deze evaluatie.

Literatuur

Adviesraad voor het basisonderwijs, speciaal onderwijs en voortgezet speciaal onderwijs (ARBO), *Voorrang aan achterstand. Advies over een integraal beleid ter voorkoming en bestrijding van onderwijsachterstand.* Zeist: Onderwijscentrum, 1988.

Beem, A.L., Y.J. Pijl & J. Scheerens, *Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma Onderwijsvoorrangsbeleid* (OVB). Den Haag: SVO, 1986.

Bergen, J.B.A.M. van, *Verantwoording constructie toetsen voor de evaluatie van het onderwijsvoorrangsbeleid.* Arnhem: CITO, 1989.

Doddema-Winsemius, H. & W.K.B. Hofstee, *Operationalisering van de effecten van het Onderwijsvoorrangsbeleid (OVB).* Den Haag: SVO, 1985.

Doddema-Winsemius, H. & M.P.C. van der Werf, *Het meten van intelligentie en sociale redzaamheid in de evaluatie van het onderwijsvoorrangsbeleid. Deelrapport 1: wetenschappelijke verantwoording; Deelrapport 2: tests en handleidingen.* Groningen: RION, 1988.

Heek, F. van, e.a., *Het verborgen talent. Milieu, schoolkeuze en schoolgeschiktheid.* Meppel: Boom, 1968.

Hoeben, W.Th.J.G. & J. Schaveling, *Evaluatie van het onderwijsvoorrangsbeleid; Haalbaar?* Groningen: RION/Intomart, 1986.

Hofstee, W.K.B., K.B. Koster & W. Meijnen, *Programmavoorstel voor evaluatieonderzoek naar de effecten van het onderwijsvoorrangsbeleid.* Groningen: Rijksuniversiteit, 1984.

Hofstee, W.K.B., *Reactie naar aanleiding van het eindrapport Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma OVB.* Groningen: Vakgroep Persoonlijkheidspsychologie (interne notitie), 1986.

Jungbluth, J., Th. Buis, G. Driessen & A. Mens, *De evaluatie van het onderwijsvoorrangsbeleid. Voorstellen voor onderzoek naar het gebruik van OVB-faciliteiten en naar de daarbij beoogde effecten.* Nijmegen: ITS, 1985.

Jungbluth, P. & W. Meijnen, *Resumerend advies over de vormgeving van de evaluatie van het onderwijsvoorrangsbeleid.* Nijmegen: ITS, 1986.

Kloprogge, J., *Uitvoeringsplan evaluatie Onderwijsvoorrangsbeleid 1987-1990.* Den Haag: SVO, 1986.
Ministerie van Onderwijs en Wetenschappen, *Beleidsplan voor het onderwijs aan groepen in achterstandssituaties. Maatregelen tot uitvoering van het Onderwijsstimuleringsbeleid.* Den Haag: Staatsuitgeverij, 1974.
Ministerie van Onderwijs en Wetenschappen, *Beleidsplan culturele minderheden in het onderwijs.* Den Haag: Staatsuitgeverij, 1981.
Ministerie van Onderwijs en Wetenschappen, *Vaststelling van een wettelijk kader voor het voeren van een samenhangend onderwijsvoorrangsbeleid (onderwijsvoorrangswet).* Den Haag: Staatsuitgeverij, 1984.
Ministerie van Onderwijs en Wetenschappen, *Onderwijsvoorrangsplan 1985-1989.* Den Haag: Staatsuitgeverij, 1985.
SVO, *Het onderwijsvoorrangsbeleid. Notitie betreffende de aanpak van de evaluatie.* Den Haag: SVO, coördinatie evaluatie onderwijsvoorrangsbeleid, 1986.
Tesser, P. & L. Mulder, *Evaluatie onderwijsvoorrangsbeleid. Notitie steekproefontwerp longitudinale onderzoeken.* Nijmegen: ITS, 1987.
Tesser, P. & G. van der Werf, Educational priority and school effectiveness. In: *Proceedings of the second meeting of the International Congress for School Effectiveness.* Lisse: Swets & Zeitlinger, 1989.
Tesser, P., M. P. C. van der Werf, L. Mulder & M. Weide, *De beginfase van het voorrangsbeleid in het basisonderwijs. Verslag van het oriënterend onderzoek van de landelijke OVB-evaluatie.* Nijmegen/Groningen: ITS/RION, 1989.
Vermeulen, H. & Werf, M. P. C. van der Werf, *De plannen van de onderwijsvoorrangsgebieden. Verslag van de analyse van de OVB gebiedsplannen.* Groningen/Nijmegen: RION/ITS, 1989.

Curricula vitae

M. P. C. van der Werf studeerde psychologie in Groningen. Sedert 1980 is zij werkzaam bij het RION, Instituut voor Onderwijsonderzoek van de Rijksuniversiteit Groningen. In 1988 promoveerde zij op "Het schoolwerkplan in het basisonderwijs". Vanaf 1986 is zij projectleider van de evaluatie van het Onderwijsvoorrangsbeleid voor het deel dat bij het RION wordt uitgevoerd.

Adres: RION, Instituut voor Onderwijsonderzoek, Postbus 1286, 9701 BG Groningen

P. Tesser studeerde psychologie in Nijmegen. Hij is sinds 1972 als onderzoeker verbonden aan het Instituut voor Toegepaste Sociale Wetenschappen (ITS) te Nijmegen. In 1986 promoveerde hij bij de Erasmus Universiteit Rotterdam op het onderwerp "Sociale herkomst en schoolloopbanen in het voortgezet onderwijs". Sinds 1986 is hij projectleider van het ITS-deel van de evaluatie van het Onderwijsvoorrangsbeleid.

Adres: Instituut voor Toegepaste Sociale Wetenschappen, Postbus 9048, 6500 KJ Nijmegen

Manuscript aanvaard 6-12-'89

Summary

Werf, M.P.C. van der & P. Tesser. 'The evaluation of the Educational Priority Policy.' *Pedagogische Studiën*, 1990, *67*, 127-138.

The Dutch Educational Priority Program aims at improving educational opportunities for disadvantaged children. The evaluation of the program started in 1987. Before the start several years were spent at the preparation of the evaluation. In this article we describe the discussions in the preparing stage about the design of and the criteria for the evaluation and the eventual decisions that were made. Also the theoretical model, the research procedure and the organizational context of the evaluation described. Finally we show some results of a preliminary study. The conclusions regard the consequences of both the careful preparations as well as the results of the preliminary study for the evaluation in the coming years.

Periodiek peilingsonderzoek: de opzet van het project PPON geïllustreerd aan de rekenpeiling einde basisonderwijs

J.M. WIJNSTRA
Instituut voor Toetsontwikkeling (Cito), Arnhem

Samenvatting

In 1986 is de voorbereiding van het project Periodieke Peiling van het Onderwijsniveau (PPON) gestart om, voorlopig beperkt tot het basisonderwijs, periodiek inhoudelijke gegevens over de stand van zaken te verzamelen. Deze gegevens moeten een basis verschaffen voor een discussie over de inhoud en het niveau van het onderwijs.

Na de inleiding worden in deze bijdrage eerst het doel en de achtergrond van het project toegelicht in relatie tot het recente beleid met betrekking tot de eindtermen. Daarna komen de organisatie, planning en werkwijze van het project in het basisonderwijs aan de orde. De werkwijze wordt in par. 4 geïllustreerd aan de eerste, in 1987 uitgevoerde peiling, de rekenpeiling einde basisonderwijs. Hierin wordt de nadruk gelegd op de bevordering van de mogelijkheden om op basis van de rapportage een discussie te voeren over de inhoud en het niveau van het onderwijs.

1 *Inleiding*

Wat moeten kinderen leren op school? En leren ze dat ook werkelijk? Over deze en soortgelijke vragen is een eindeloze discussie mogelijk, juist door het gebrek aan inhoudelijke gegevens over het onderwijs. Om aan de discussie over de inhoud en het niveau van het onderwijs een meer rationele basis te verschaffen is in 1986 in opdracht van de Minister van Onderwijs en Wetenschappen de voorbereiding van het project Periodieke Peiling van het Onderwijsniveau (PPON) van start gegaan om – voorlopig beperkt tot het basisonderwijs – periodiek inhoudelijke gegevens over de stand van zaken te verzamelen.

Hoewel er ook relaties zijn met de voorziene eindtermen basisonderwijs, is het project daarvan geen uitvloeisel. De initiatieven tot het opzetten van periodiek peilingsonderzoek dateren al van het eind van de jaren zeventig. Onder verwijzing naar een advies uit de SVO-nota Opstaan tegen het Zittenblijven (Doornbos, 1969) heeft de Innovatiecommissie Basisonderwijs (ICB) reeds in 1978 gepleit voor een periodiek onderzoek naar de kwalitatieve ontwikkeling van het basisonderwijs, waarbij de nadruk zou moeten liggen op de inventarisatie van het leer- en vormingsaanbod (ICB, 1978). In de beleidsnotitie bij dit innovatieplan heeft de toenmalige minister positief gereageerd op deze aanbeveling. Hoewel er reeds eind 1979 een technisch vooroverleg over dit onderwerp werd gestart met de SLO, de SVO en het Cito om de mogelijkheden te verkennen, heeft het tot juni 1985 geduurd voor de beleidsnotitie verscheen die ten grondslag ligt aan de opdracht om peilingsonderzoek in het onderwijs te gaan uitvoeren. In die tussentijd verscheen de nota Kwaliteit van het Onderwijs (1981), terwijl in verschillende memories van toelichting bij het onderwijshoofdstuk uit de rijksbegroting aandacht werd besteed aan de kwaliteitsproblematiek. Aan de verstrekking van de opdracht ging verder een door de Stichting Centrum voor Onderwijsonderzoek (SCO) te Amsterdam uitgevoerde voorstudie op het gebied van het moedertaalonderwijs vooraf waarin de haalbaarheid van periodiek peilingsonderzoek werd verkend (zie Wesdorp, 1985; Wesdorp, Van den Bergh, Bos, Hoeksma, Oostdam, Scheerens & Triesscheijn, 1986).

In deze bijdrage wordt de opzet van het project PPON in het basisonderwijs geschetst[1]. Na een beschrijving van het doel en de achtergrond van het project worden de organisatie, planning en werkwijze toegelicht. Bij wijze van illustratie wordt de bijdrage afgerond met een beschrijving van de eerste uitgevoerde peiling, de rekenpeiling einde basisonderwijs. Daarbij wordt de nadruk gelegd op de bevordering van de mogelijkheden om op basis van de rapportage een discussie te

voeren over de inhoud en het niveau van het onderwijs.

2 *Doel en achtergrond van PPON*

In de beleidsnotitie van juni 1985 die aan de opdracht ten grondslag ligt, wordt gesignaleerd dat de aandacht voor het vraagstuk van de kwaliteit van het onderwijs de laatste jaren sterk is toegenomen. Tegelijkertijd wordt vastgesteld 'dat de onderwijskundige en maatschappelijke discussies over het niveau en de kwaliteit van het onderwijs worden bemoeilijkt door het veelal ontbreken van empirische gegevens over de resultaten van het onderwijs'. Om in deze leemte te voorzien is daarom aan de Stuurgroep PPON, waarin aanvankelijk het Instituut voor Onderzoek van het Onderwijs (SVO) en het Instituut voor Toetsontwikkeling (Cito) samenwerkten, opdracht gegeven peilingsonderzoek te gaan uitvoeren, te beginnen in het basisonderwijs.

In de aangehaalde beleidsnotitie wordt het algemene doel van peilingsonderzoek omschreven als 'systematisch bij te dragen tot het verkrijgen van een beeld van het leeraanbod en de effecten van het onderwijs' met als voornaamste functies:

- het bieden van een referentiekader voor de school voor de beoordeling van de eigen doelstellingen en resultaten;
- het leveren van indicaties voor het verrichten van nadere (diepte-)studies en evaluatie-onderzoek die van belang kunnen zijn voor de kwaliteitsbeoordeling en kwaliteitsbevordering van het onderwijs;
- het bijdragen aan de ontwikkeling van toetsen voor het gebruik door scholen, die een vergelijking met landelijke trends mogelijk maken;
- het bieden van een (minimale) empirische basis voor de meer algemene maatschappelijke discussie over kwaliteit en niveau van het onderwijs.

PPON moet in principe het gehele curriculum van het basisonderwijs bestrijken. De peilingsonderzoeken richten zich voorlopig primair op de stand van zaken aan het einde van het basisonderwijs, met uitzondering van Nederlandse taal en rekenen waarvoor ook peilingsonderzoeken halverwege het basisonderwijs worden uitgevoerd. Peilingsonderzoek is in principe steekproefonderzoek, dit in tegenstelling tot bijvoorbeeld de bestandstellingen van het Centraal Bureau voor de Statistiek en het Ministerie van Onderwijs en Wetenschappen. Het is de taak van het project PPON op basis van gegevensverzameling bij steekproeven van scholen en leerlingen per leer- en vormingsgebied, periodiek herhaald, een algemeen inzicht te verschaffen in het actuele leer- en vormingsaanbod, de onderwijsleerresultaten en de relaties daartussen. Daarbij dient ook aandacht te worden geschonken aan relaties van deze variabelen met andere relevante kenmerken van scholen en leerlingen, zodat in de rapportage relevante subgroepen afzonderlijk kunnen worden beschreven. Het is echter uitdrukkelijk niet de bedoeling – en in de gekozen opzet ook niet goed mogelijk – uitspraken over individuele scholen of leerlingen te doen. Wel levert PPON een referentiekader voor de zelfevaluatie van scholen en wordt nagegaan in hoeverre de peilingsinstrumenten voor zelfevaluatie kunnen worden ingezet.

Gezien het tijdstip waarop de beleidsnotitie tot stand kwam, kon in de weergegeven opsomming van functies nog niet worden verwezen naar de (toekomstige) eindtermen basisonderwijs, waarmee uiteraard ook relaties zullen gaan bestaan als het wetsvoorstel dienaangaande wordt aangenomen. Volgens de Memorie van Toelichting bij dit wetsvoorstel zijn eindtermen 'beschrijvingen van kwaliteiten van leerlingen op het gebied van kennis, inzicht en vaardigheden die de school bij haar onderwijsactiviteiten in elk geval als doelstelling moet hanteren. Eindtermen maken deel uit van de kwaliteitseisen die aan het basisonderwijs (...) worden gesteld. In die zin zijn eindtermen richtlijnen voor scholen voor de na te streven resultaten van leerlingen'. Gezien deze formulering zijn eindtermen op te vatten als een normstelling in het (cyclische) proces van kwaliteitsbeheersing (vgl. de nota Kwaliteit van het Onderwijs, 1981) en worden deze ook van belang voor de inhoud van de peilingsonderzoeken.

Door dit gewijzigde kader, dat enkele jaren geleden – althans in het basisonderwijs – nog moeilijk voorstelbaar was (vgl. Wijnstra, 1982) en zijn oorsprong vindt in het WRR-rapport over de basisvorming (WRR, 1986), wordt (impliciet) de functie van PPON ook verruimd, in die zin dat het project (mede) een functie krijgt in de landelijke kwaliteitsbe-

heersing van het onderwijs. In de terminologie van de aangehaalde nota vormen de peilingsonderzoeken een instrument om op het niveau van het onderwijssysteem de mate van normrealisering waar te nemen (de tweede fase in het proces van kwaliteitsbeheersing).

Dit betekent dat in de toekomst de eindtermen in elk geval in de peilingsinstrumenten vertegenwoordigd moeten zijn om een afweging en waardering van de bevindingen in relatie tot de normstelling (de derde fase in het proces van kwaliteitsbeheersing) mogelijk te maken. Hierdoor vindt overigens tegelijkertijd een concretisering van de eindtermen plaats, zodat ook een oordeel over de haalbaarheid daarvan mogelijk is. Men kan zich in dit verband zelfs afvragen of het formuleren van eindtermen zinvol is als deze niet operationeel gedefinieerd worden. Peilingsonderzoek behoeft zich echter niet te beperken tot de eindtermen. Als dat het geval was, zou een discussie over ontwikkelingen in het onderwijs nodeloos worden gefrustreerd. Tegelijkertijd wordt ook het gevaar van een conserverend effect voorkomen. Juist met het oog op een discussie over nieuwe ontwikkelingen is het van belang ook zaken die nog geen gemeengoed zijn, maar wel al een behoorlijke ingang hebben gevonden, in de peilingen op te nemen. Dit principe is ook het uitgangspunt bij de voorbereiding van de peilingsonderzoeken in de periode dat er nog geen officiële eindtermen zijn.

PPON levert inhoudelijke beschrijvingen van de stand van zaken op basis waarvan een gefundeerde discussie over de inhoud en het niveau van het onderwijs gevoerd kan worden en eventueel maatregelen kunnen worden getroffen ter handhaving, respectievelijk bijstelling van de normen, zoals de vierde fase in het proces van kwaliteitsbeheersing wordt genoemd in de nota Kwaliteit van het Onderwijs (1981). Het waarderen van de bevindingen en het doen van aanbevelingen maken echter geen deel uit van de opdracht. PPON rekent het echter wel tot zijn taak de discussie over de peilingsuitkomsten te bevorderen.

3 *Organisatie, planning en werkwijze*

3.1 *Organisatie*

Oorspronkelijk is de opdracht tot het uitvoeren van het peilingsonderzoek verleend aan een samenwerkingsverband van SVO en Cito, die daartoe de Stuurgroep PPON hadden gevormd. Voorzien was dat het Cito de periodiek herhaalde peilingen zou uitvoeren en dat via SVO dieptestudies naar aanleiding van bevindingen uit de peilingen zouden worden uitgevoerd. Voor het geheel zou één budget beschikbaar zijn dat door SVO zou worden beheerd. Deze taakverdeling bestaat nog steeds, de vorm van het samenwerkingsverband en het financieel beheer zijn echter herzien. Enerzijds gaf de invoering van de Wet op de Onderwijsverzorging (WOV) hiertoe aanleiding, anderzijds bleek de gezamenlijke verantwoordelijkheid van SVO en Cito voor zowel de programmering als de uitvoering problemen op te leveren. De Cito-activiteiten worden nu op basis van artikel 73 van de WOV gefinancierd (1.6 miljoen op jaarbasis), terwijl SVO en Cito een coördinatiegroep hebben gevormd waarin afstemmingsoverleg plaatsvindt. SVO en Cito hebben daarnaast de gezamenlijke adviesgroep met vertegenwoordigers van de Vereniging Samenwerkende Landelijke Pedagogische Centra (VSLPC), de Vereniging Werkverband van Plaatselijke en Regionale Onderwijsbegeleidingsdiensten (WPRO), het Instituut voor Leerplanontwikkeling (SLO) en waarnemers van het Ministerie van Onderwijs en Wetenschappen en de Inspectie Basis- en Speciaal Onderwijs met een nagenoeg ongewijzigde taakomschrijving gehandhaafd. De taken van deze adviesgroep betreffen met name advisering over de gebruikswaarde van de projectopbrengst en de doorwerking daarvan in het onderwijs.

3.2 *Planning*

In een periode van ongeveer vijf jaar moeten alle leer- en vormingsgebieden uit de Wet op het Basisonderwijs (WOB) in peilingsonderzoeken aan de orde komen. Zeker ook in het licht van de eindtermen moet het als een gelukkige beslissing worden gezien dat het gehele curriculum moet worden bestreken. Daarmee wordt voorkomen dat eenzijdige accenten worden gelegd die ongewenste neveneffecten kunnen hebben. Na een evaluatie van de

eerste cyclus, waarover gerapporteerd wordt aan de Centrale Commissie voor Onderwijsoverleg (CCOO) en de Tweede Kamer, begint daarna in principe de volgende cyclus van peilingsonderzoeken, zodat ook vergelijkingen over de tijd mogelijk worden. Voor de eerste cyclus geldt op dit moment het tijdschema dat in Tabel 1 is weergegeven.

Tabel 1 *Tijdschema voor de eerste cyclus van peilingsonderzoeken in het basisonderwijs*

Rekenen	1987
Nederlandse taal	1988/1989
Sociale wereldoriëntatie	1990
Natuuronderwijs en gezondheidseducatie	1990/1991
Verkeerseducatie	1991
Engelse taal	1991
Bewegingsonderwijs	1987/ ?
Kunstzinnige vorming	1987/1988/ ?

De rekenpeilingen medio en einde basisonderwijs zijn in 1987 uitgevoerd en de rapportage daarover is inmiddels beschikbaar (Wijnstra, 1988, 1990; Verhelst & Eggen, 1989). In 1988 heeft de gegevensverzameling voor de taalpeiling einde basisonderwijs plaatsgevonden (rapportage voorjaar 1990), terwijl de mediopeiling op dit gebied in het najaar van 1989 is uitgevoerd. Bij de voorbereiding en verwerking van de taalpeilingen wordt samengewerkt met de SCO te Amsterdam.

Zoals eerder aangegeven worden op de overige leer- en vormingsgebieden alleen peilingsonderzoeken aan het ejnde van het basisonderwijs uitgevoerd. Wereldoriëntatie is in de planning in twee grote onderdelen gesplitst: sociale wereldoriëntatie (alle kennisgebieden uit de WOB met uitzondering van natuur) en natuuronderwijs, met inbegrip van gezondheidseducatie. De voorbereidingen voor wereldoriëntatie zijn al in de loop van 1986 gestart, in samenwerking met het Citoproject Itembank Wereldoriëntatie. Voor wereldoriëntatie is relatief veel voorbereidingstijd uitgetrokken, omdat de leerplanontwikkeling en toetsconstructie op dit gebied, in vergelijking met andere leer- enrvormingsgebieden, nog niet erg ver ontwikkeld zijn.

Ter voorbereiding van peilingen op het gebied van de kunstzinnige vorming (de expressie-activiteiten uit de WOB) en het bewegingsonderwijs zijn in de jaren 1987 en 1988 in samenwerking met de SLO inventarisaties van het leer- en vormingsaanbod uitgevoerd met behulp van vragenlijsten (Brink, Wijnstra & Janssens, 1989a, 1989b; Zwarts & Wijnstra, 1989). In 1989 is de voorstudie naar de vorm en inhoud van peilingsonderzoeken op deze leer- en vormingsgebieden afgerond, wederom in samenwerking met de SLO. Gelet op het karakter van het onderwijs op deze gebieden diende nader bepaald te worden welke doelstellingen voor meting van kennis en vaardigheden op leerlingniveau in aanmerking komen en welke doelstellingen via een meting van het onderwijsaanbod moeten worden gepeild. Waarschijnlijk zal de eerste peiling in 1992 plaatsvinden. Welk onderdeel dit betreft, is nog niet vastgesteld. Dit verklaart de vraagtekens in Tabel 1.

3.3 *Werkwijze*

Naast instrumenten om achtergrondkenmerken van scholen en leerlingen te inventariseren, zijn – gelet op de doelstellingen van PPON – twee soorten, gedeeltelijk complementaire instrumenten nodig: instrumenten om het leer- en vormingsaanbod te inventariseren en instrumenten om vaardigheden (met inbegrip van kennis en inzicht) bij leerlingen te meten. In deze paragraaf wordt een aantal uitgangspunten voor de instrumentontwikkeling geformuleerd.

De instrumentontwikkeling voor een leer- en vormingsgebied start met het opstellen van een structurele beschrijving van het betreffende gebied in de vorm van een geordende lijst van leer- en vormingsdoelen op basis van aanwezige expertise, beschikbare leergangen, handboeken, enzovoort. Na een commentaarfase wordt de domeinbeschrijving vastgesteld, waarna deze dienst gaat doen als uitgangspunt voor de feitelijke instrumentontwikkeling. Zoals ook aangegeven in de vorige paragraaf is het mogelijk om in de peilingsonderzoeken elementen op te nemen die in het onderwijs (nog) geen gemeengoed zijn. Als op voorhand de inhoud van een leer- en vormingsgebied zou worden beperkt tot wat algemeen gangbaar is, of – in de toekomst – de wettelijke eindtermen basisonderwijs, snijdt dat de mogelijkheid af wellicht toch interessante of belangrijke aspecten voor de ontwikkeling van het onderwijs in de discussie over de bevindingen te betrekken. Dit impliceert ook dat de inhoud van het peilingsinstru-

mentarium niet gelegitimeerd behoeft te worden in de zin van de verantwoording van beperkende keuzen, zoals dat voor leerplan- en leergangontwikkeling geldt.

Dezelfde domeinstructurering leent zich voor de ontwikkeling van geheel verschillende soorten toetsen, afhankelijk van het gebruiksdoel. Als het om een algemene niveau-indicatie gaat, kan met een betrekkelijk korte, maar inhoudelijk gevarieerde toets worden volstaan. De resultaten op een dergelijke toets lenen zich echter niet goed voor verdere inhoudelijke uitsplitsing. In het ontwerp van PPON heeft voorop gestaan dat de leer- en vormingsgebieden ruim gedekt moeten worden en dat de resultaten een inhoudelijke discussie over het onderwijs mogelijk moeten maken, bijvoorbeeld in het kader van de ontwikkeling van eindtermen. Dit heeft tot het uitgangspunt geleid dat de instrumenten een beschrijving van de vaardigheden van de leerlingen mogelijk moeten maken op nauwkeurig omschreven, didactisch betekenisvolle eenheden van beperkte omvang.

Dit niveau van detaillering komt ongeveer overeen met wat Bock, Mislevy en Woodson (1982) 'indivisible curricular elements' noemen en in dit opzicht komt de opzet van de PPON overeen met die van het California Assessment Program (CAP; vgl. Pandey, 1986). Dezelfde principes zijn ook terug te vinden in het proefonderzoek van Wesdorp et al. (1986). Een dergelijke opzet leidt tot een vrij groot aantal rapportage-eenheden per leer- en vormingsgebied. Dat is voor een discussie over de inhoud en het niveau van het onderwijs of voor de planning van ontwikkelingswerk eerder een voor- dan een nadeel. Uiteraard zullen de vaardigheden van de leerlingen op verwante rapportage-eenheden gecorreleerd zijn. Gelet op het doel van het onderzoek is dit echter van secundair belang.

Een dergelijke modulaire opbouw heeft daarnaast diverse andere voordelen, vooral met het oog op de vergelijking over de tijd. Bij kleine eenheden is de kans het grootst dat veranderingen in het curriculum een uniform effect hebben op alle opgaven van een module, zodat veranderingen in de vaardigheid van de leerlingen goed te interpreteren zijn. Als uitgegaan zou worden van meer omvattende curriculumeenheden, zouden accentwijzigingen in het curriculum nagenoeg onoverkomelijke problemen opleveren bij de vergelijking over de tijd. Een verder voordeel is dat nieuwe aspecten op een gebied gemakkelijk als modulen aan het instrumentarium kunnen worden toegevoegd zonder dat de vergelijkbaarheid van scores op andere modulen met eerdere peilingen wordt aangetast (verouderde schalen kunnen ook zonder bezwaar worden weggelaten). Betrekkelijk kleine, inhoudelijk homogene eenheden bieden ten slotte – voorzover dat van toepassing is – een betere mogelijkheid itemresponsmodellen toe te passen dan meer omvattende eenheden. In tegenstelling tot de klassieke testtheorie is het met behulp van itemresponsmodellen mogelijk constante schalen te definiëren, onafhankelijk van de vaardigheid van de leerlingen. Daardoor is het ook mogelijk in een later stadium nieuwe opgaven aan de schaal toe te voegen, zodat in een volgende peilingsronde met gedeeltelijk nieuwe opgaven de verdeling van de vaardigheid in de populatie op een constante schaal kan worden weergegeven. Vervanging van opgaven is noodzakelijk, omdat een deel van de opgaven als illustratie in de rapportage moet worden gebruikt.

De benodigde opgaven voor de peilingsinstrumenten zullen in het algemeen worden geconstrueerd door commissies van ervaren leraren uit het basisonderwijs onder leiding van een projectmedewerker. De vorm van de opgaven zal afhankelijk zijn van de te meten vaardigheid, waarbij het uitgangspunt is dat het domein zo goed mogelijk gedekt moet worden. In dit opzicht worden op voorhand geen beperkingen aangebracht. Als een vaardigheid beter met open vragen kan worden gemeten dan met meerkeuzevragen, wordt de voorkeur gegeven aan open vragen. Als groepsgewijze, schriftelijke afname van de opgaven geen adequate mogelijkheid is, zal gebruik worden gemaakt van individuele afnameprocedures.

Complementair aan de instrumenten voor het meten van de onderwijsleerresultaten dienen instrumenten voor de inventarisatie van het leer- en vormingsaanbod te worden samengesteld. Deze zullen over het algemeen bestaan uit schriftelijke vragenlijsten voor de groeps- en vakleraren en/of de directeur van de school, maar ook andere instrumentaties zijn denkbaar. De inhoud en vorm van deze instrumenten zal sterk afhankelijk zijn van het leer- en vormingsgebied. Waar dat van toepassing is, zal ook aandacht worden

besteed aan aspecten die niet direct de bedoeling hebben tot bepaalde leerresultaten te leiden ('expressive objectives' in de terminologie van Eisner, 1969) en die niet in de toetsen vertegenwoordigd zijn.

4 *Rekenpeiling einde basisonderwijs: een illustratie*

4.1 *Inleiding*

In 1987 zijn de hiervoor beschreven uitgangspunten en werkwijzen voor de eerste keer toegepast in de rekenpeilingen. In het voorjaar is de peiling einde basisonderwijs uitgevoerd en in het najaar de peiling medio basisonderwijs. De peilingen werden op afzonderlijke steekproeven van scholen uitgevoerd. In deze paragraaf worden als illustratie van het voorgaande enkele aspecten nader belicht op basis van de rekenpeiling einde basisonderwijs. Daarbij wordt de nadruk gelegd op de bevordering van de mogelijkheden om naar aanleiding van de bevindingen een discussie te voeren over de inhoud en het niveau van het onderwijs. Voor (methodologische) details zij verwezen naar Bokhove, Brink en Janssen (1988a), Eggen (1989), Verhelst (1989), Verhelst en Eggen (1989) en Wijnstra (1988, 1989).

4.2 *Dataverzameling en -verwerking*

Afgezien van de instrumenten om gegevens over achtergrondkenmerken van scholen en leerlingen te verzamelen, werden de instrumenten gebaseerd op een indeling van het rekendomein in 27 onderwerpen, verdeeld over tien hoofdonderwerpen (zie Tabel 2). Aan deze indeling ligt een kruising van inhouds- en gedragscategorieën ten grondslag.

In samenwerking met een tweetal commissies van ervaren leraren uit het basisonderwijs zijn bijna 500 opgaven geconstrueerd: 32 voor het onderwerp basisoperaties en zestien à twintig voor elk van de overige onderwerpen, voor het merendeel in open vraagvorm. Aan deze vorm werd de voorkeur gegeven om tot een zo volledig mogelijke dekking te komen en om zo dicht mogelijk aan te sluiten bij de praktijk van het onderwijs. Alleen wanneer de meerkeuzevorm duidelijke voordelen had, bijvoorbeeld bij schattingsopgaven, werd deze vorm gekozen. In psychometrisch opzicht heeft de open vraagvorm als voordeel dat geen rekening behoeft te worden gehouden met effecten van raden.

De toetsen werden onder leiding van toetsassistenten op ruim 300 scholen gemaakt. Om tijdsredenen en om het clustereffect te beperken maakte iedere leerling slechts een deel van de toetsen. Feitelijk werden op deze wijze zes (equivalente) substeekproeven gecreëerd met ieder 1200 à 1300 leerlingen. In aanvulling op de klassikale toetsafnames werden per klas, afhankelijk van het aantal leerlingen, twee of drie leerlingen systematisch getrokken die in een individuele sessie drie tot vijf van de eerder gemaakte opgaven over één onderwerp nog eens opnieuw kregen voorgelegd met de vraag aan te geven welke procedure bij de oplossing was gevolgd. Dit betrof in totaal dertig opgaven over acht onderwerpen, met name hoofdreken- en toepassingsonderwerpen. De procedures werden door de toetsassistenten volgens een geprecodeerd categorieënsysteem ingedeeld of uitgeschreven, als de werkwijze niet direct was onder te brengen.

Alle 27 onderwerpen konden met toepassing van een itemresponsmodel succesvol worden geschaald waarbij minder dan 5% van de opgaven buiten beschouwing moest blijven, omdat deze onvoldoende in het model pasten. Onder aanname van een bij benadering standaardnormale verdeling van de vaardigheden van de leerlingen (door de toegepaste transformaties) is bij alle schalen het landelijk gemiddelde van de leerlingen in jaargroep 8 op 250 gesteld en de standaarddeviatie op 50. Door de toepassing van een itemresponsmodel kunnen de moeilijkheidsgraad van de opgaven en de vaardigheid van de leerlingen op dezelfde schaal worden afgebeeld (zie voor een voorbeeld Figuur 1). In de gekozen opzet voor de rapportage geeft de schaalwaarde van een opgave aan dat een leerling met een overeenkomstige vaardigheidsscore een kans van 80% heeft een dergelijke opgave goed te maken. Bij lager op de schaal geplaatste opgaven heeft dezelfde leerling een grotere succeskans; bij moeilijker, hoger op de schaal geplaatste opgaven is die kans uiteraard kleiner. Naast de schaal zijn de percentielen vermeld (P..), dat wil zeggen de percentages leerlingen die een score onder of gelijk aan de betreffende schaalwaarde hebben behaald. In aanvulling op deze gegevens kan aan de hand van de gegevens uit de individuele afnames inzicht worden gegeven in de gehan-

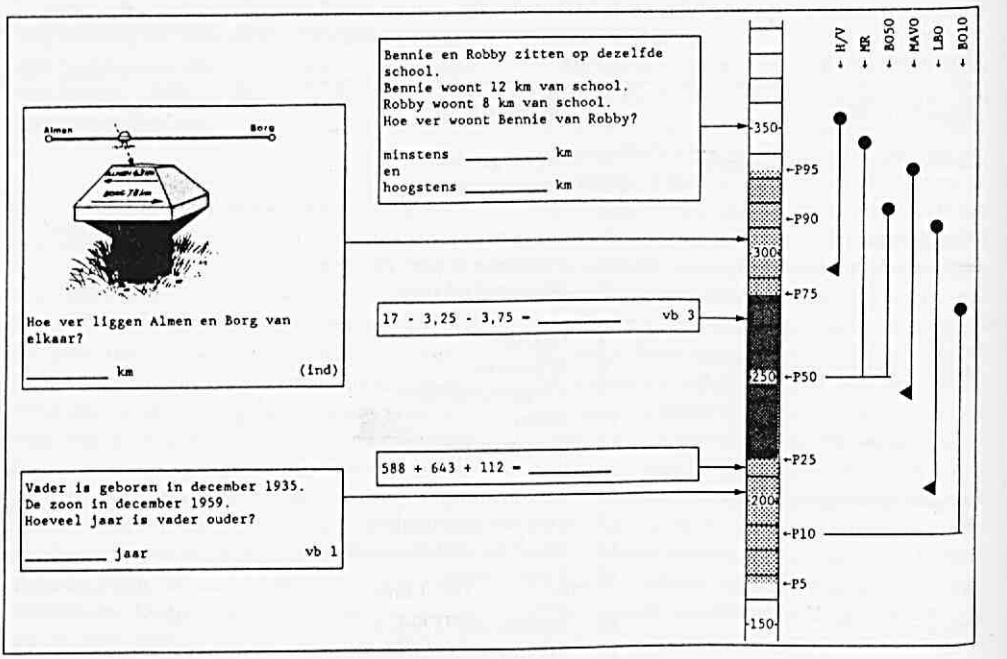
Tabel 2 *Rekenonderwerpen peiling einde basisonderwijs*

Hoofdonderwerp	Onderwerp	
Basisoperaties	1	Basisoperaties
Getallen: basiskennis en begrip	2	Gehele getallen
	3	Kommagetallen
Hoofdrekenen (gehele en kommagetallen)	4	Optellen en aftrekken
	5	Vermenigvuldigen en delen
	6	Schattend rekenen
Cijferen (gehele en kommagetallen)	7	Optellen
	8	Aftrekken
	9	Vermenigvuldigen
	10	Delen
	11	Toepassingen
Breuken/gemengde getallen	12	Basiskennis en begrip
	13	Optellen en aftrekken
	14	Vermenigvuldigen en delen
	15	Toepassingen
Procenten	16	Basiskennis en begrip
	17	Handig procentrekenen
	18	Toepassingen
Verhoudingen	19	Basiskennis en begrip
	20	Toepassingen
Meten (lengte, omtrek, oppervlakte, inhoud, gewicht)	21	Basiskennis en begrip
	22	Tellen/aflezen van het aantal maateenheden
	23	Berekenen van het aantal maateenheden
	24	Maatsystemen
	25	Toepassingen
Kalender en klok	26	Toepassingen
Geld	27	Toepassingen

teerde oplossingswijzen. De in Figuur 1 met (ind) gemerkte opgave is een voorbeeld van een opgave die ook individueel werd afgenomen.

Om zicht te krijgen op de omvang en de inhoud van het rekenonderwijs vulden de groepsleraren van de jaargroepen 6, 7 en 8 een vragenlijst in. Naast enkele vragen over de methode, differentiatie en de beschikbare tijd voor het rekenonderwijs, stond hierin de (relatieve) verdeling van de tijd over de in de domeinbeschrijving onderscheiden (hoofd)onderwerpen centraal. De groepsleraar van jaargroep 8 kreeg verder per onderwerp drie a priori naar moeilijksheidsgraad geordende voorbeeldopgaven voorgelegd met het verzoek aan te geven tot hoe ver de behandeling van de rekenstof was gevorderd voor de gemiddelde leerling (in Figuur 1 zijn twee van dergelijke met (vb) gemerkte opgaven afgebeeld). Voor de deelnemende leerlingen moest op de leerlingenlijst tevens worden vermeld welke leerlingen eventueel de stof die nodig is om de gemakkelijkste voorbeeldopgave op te lossen, nog niet hadden gehad.

In de rapportage worden de aanbod- en vaardigheidsgegevens met elkaar in verband gebracht, terwijl ook relaties met achtergrondkenmerken van scholen en leerlingen worden geëxploreerd.



Figuur 1 *Schaal 4: Optellen en aftrekken (hoofdrekenen)*

4.3 *Gerealiseerd en gewenst niveau*

In de rapportage zoals weergegeven in Figuur 1, wordt het gerealiseerde niveau van leerlingen aan het einde van het basisonderwijs beschreven. Of dit ook een acceptabel niveau is, is een geheel andere vraag, die niet door PPON kan worden beantwoord. Omdat dit de eerste peiling is, zijn ook nog geen vergelijkingen met eerdere peilingen mogelijk. Ter bevordering van de discussie is besloten enige indicaties te verzamelen over het gewenste niveau van enkele groepen leerlingen. Daartoe zijn in het voorjaar van 1988, toen de schaalconstructie voorlopig was afgerond, enquêtes gehouden onder groepsleraren van jaargroep 8, ouderleden van medezeggenschapsraden en docenten uit het LBO, MAVO en HAVO/VWO.

De enquêtes bestonden uit afbeeldingen van de schalen over de onderwerpen 2–27 met de illustratieve opgaven, maar zonder gegevens over de peilingsresultaten. De ouderleden van de medezeggenschapsraden en de leraren van de basisscholen werd een oordeel gevraagd over het gewenste niveau van de gemiddelde leerling op ieder van de schalen. De groepsleraren van de basisscholen werd tevens gevraagd aan te geven welk niveau tenminste door 90% van de leerlingen zou moeten worden gehaald. De docenten uit (de brugklas van) het voortgezet onderwijs werd gevraagd aan te geven welk niveau voor hun schooltype aan het begin van het voortgezet onderwijs gewenst werd geacht.

Parallel aan de vaardigheidsschaal in Figuur 1 wordt het door de respondenten gewenste niveau en het verschil met het (geschatte) feitelijk gerealiseerde niveau aangegeven. Over het algemeen blijken aanzienlijke discrepanties. Bij de afgebeelde schaal leggen de groepsleraren van jaargroep 8 bijvoorbeeld gemiddeld het minimaal gewenste niveau voor 90% van de leerlingen (in Figuur 1 aangeduid als BO10) op 278. Dit wordt aangegeven door het bolletje. Het verschil met het niveau dat 90% van de leerlingen werkelijk bereikt (percentiel 10) wordt door de verticale lijn aangegeven en bedraagt bijna 90 punten op de vaardigheidsschaal, bijna twee standaarddeviaties. Het gewenste niveau voor deze groep leerlingen ligt hoger dan het feitelijke landelijke gemiddelde. De groepsleraren plaatsen het gewenste niveau van de gemiddelde leerling (BO50) op 317, de ouderleden van de mede-

zeggenschapsraden (MR) komen uit op 344, een niveau dat door nog geen 5% van de leerlingen wordt gehaald. Het gewenste niveau van de gemiddelde leerling wordt in de figuur door een lijn verbonden met percentiel 50. Gelet op het gerealiseerde niveau plaatsen ook de docenten uit het voortgezet onderwijs de grenzen hoog. De LBO-docenten leggen (gemiddeld) het gewenste niveau op 310. De MAVO-docenten komen uit op 332 en de HAVO/VWO-docenten (H/V) op 353. Deze niveaus worden door een lijn verbonden met de geschatte gemiddelden van deze groepen leerlingen op basis van gegevens uit de Eindtoets Basisonderwijs. Het vlaggetje aan de lijn wijst het betreffende schaalpunt ongeveer aan.

Om in de rapportage een tweede voorzet te geven voor een discussie over de inhoud en het niveau van het rekenonderwijs zijn een drietal deskundigen op dit gebied uitgenodigd een commentaar te schrijven op de uitkomsten van de peiling einde basisonderwijs. In de commentaren zou het vooral moeten gaan om de vraag wat het rekenonderwijs moet inhouden en welke resultaten dat moet opleveren, waarbij ook aanbevelingen werden gevraagd met betrekking tot de didactiek, ontwikkelingswerk, opleiding, begeleiding en onderzoek. Zonder verschuivingen in de tijdsverdeling binnen het rekenonderwijs uit te sluiten, werd de commentatoren verzocht uit te gaan van de randvoorwaarde dat voor het totaal van het rekenonderwijs in de toekomst (gemiddeld) niet meer tijd beschikbaar is dan in de tegenwoordige situatie (bijna vijf uur per week in de bovenbouw).

De commentaren zijn integraal opgenomen in een afzonderlijk deel van het rapport (zie Wijnstra, 1988). In het eerste hoofdstuk komt Huitema, medewerker van de Onderwijsbegeleidingsdienst in het Gewest Arnhem aan het woord. Hij is voorzitter van de Centrale Werkgroep Rekenen/Wiskunde-Onderwijs en tevens auteur van de rekenmethode De Wereld in getallen en het remediërend rekenpakket Remelka. Teunissen, universitair docent aan de Vakgroep Onderwijskunde van de Rijksuniversiteit in Utrecht, neemt het tweede hoofdstuk voor zijn rekening. Hij is op het gebied van het rekenonderwijs vooral bekend als auteur van zowel de eerste als de vernieuwde versie van de methode Operatoir Rekenen. In het laatste hoofdstuk geeft Treffers, hoogleraar aan de Vakgroep Onderzoek Wiskundeonderwijs en Onderwijs Computercentrum (OW & OC) van de Rijksuniversiteit in Utrecht zijn visie. Hij werkte eerder mee aan het Wiskobas-project en promoveerde op het proefschrift Wiskobas Doelgericht (1978).

4.4 *Presentatie en verspreiding van de bevindingen*

Mede voorbereid door een serie artikelen van Bokhove en Janssen (1987), respectievelijk Bokhove, Brink en Janssen (1988a, 1988b) zijn de bevindingen uit de rekenpeiling einde basisonderwijs en de commentaren op 2, 3 en 4 november 1988 uitvoerig gepresenteerd op de jaarlijkse najaarsconferentie van de Nederlandse Vereniging tot Ontwikkeling van het Reken-Wiskundeonderwijs (NVORWO), waarbij in het programma ook nadrukkelijk de relatie werd gelegd met de eindtermen en met name de zogenaamde beredeneerde eindtermen voor rekenen/wiskunde (Treffers, 1988). Als vervolg op de conferentie zijn verdere artikelen gepubliceerd in het Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs en Willem Bartjens (Bokhove & Janssen, 1988c, 1989a-e) en er zullen nog enkele volgen. Op deze manier kunnen ontwikkelaars, begeleiders en opleiders op het gebied van het rekenonderwijs op een succesvolle manier worden bereikt. De betreffende instituten hebben verder het volledige rapport over de rekenpeilingen ontvangen.

Alle scholen voor basis- en speciaal onderwijs hebben een brochure ontvangen met een beknopte schets van de rekenpeilingen. Met name voor een discussie over het schoolwerkplan van een individuele school blijft het echter een probleem dat de peilingsonderzoeken zich op macro-niveau bewegen. Hierdoor is het moeilijk de verbinding naar de eigen situatie van een individuele school te leggen. De mogelijkheid om de situatie van de eigen school te kunnen vergelijken met landelijke of regionale gegevens is in verschillende 'state assessment programs' in de Verenigde Staten zeer effectief gebleken om op lokaal niveau het doel en het functioneren van het onderwijs ter discussie te stellen en – indien gewenst – bij te stellen (met name de staat Minnesota is een goed voorbeeld; vgl. Wijnstra, 1982 en Richards, 1988).

Gezien de opzet van PPON is het in het kader van de gewone gegevensverzameling niet

mogelijk om op schoolniveau te rapporteren. Nog afgezien van de zinvolheid van een dergelijke rapportage in steekproefonderzoek, verhindert de toegepaste vorm van steekproeftrekking binnen de scholen dit, omdat niet alle leerlingen dezelfde opgaven hebben gemaakt. In de politieke besluitvorming over het peilingsonderzoek is deze mogelijkheid ook uitgesloten. Scholen dienen op vrijwillige basis deel te nemen, zonder in de rapportage herkenbaar te zijn om ongewenste sancties te voorkomen. Wel dienen de peilingsonderzoeken een kader te scheppen voor zelfevaluatie en dient te worden bijgedragen aan de ontwikkeling van toetsen die scholen in dit verband kunnen toepassen (vgl. par. 2).

In verband met een eventueel toekomstig gebruik van de PPON-instrumenten in het kader van de zelfevaluatie is in 1988 een proef genomen met een experimentele schoolrapportageservice. De scholen die in 1987 aan de rekenpeilingen hadden deelgenomen, werden in de gelegenheid gesteld het instrumentarium voor de rekenpeiling einde basisonderwijs in gecomprimeerde vorm af te nemen bij de leerlingen die in 1988 de basisschool zouden verlaten. Ruim een derde van de ongeveer 600 scholen heeft hieraan deelgenomen. Begin 1989 hebben zij een zodanige rapportage over combinaties van de oorspronkelijke schalen ontvangen dat een vergelijking mogelijk is met landelijke gegevens en met de categorie van scholen met een vergelijkbare samenstelling van de schoolbevolking. Uit de evaluatie is gebleken dat deze service door de scholen positief wordt gewaardeerd. Gezien het succesvolle verloop van deze proef heeft het Cito het voornemen deze service verder te ontwikkelen en algemeen beschikbaar te stellen, ook met uitbreiding naar andere leer- en vormingsgebieden. Deze activiteiten zullen in een afzonderlijk serviceproject worden ondergebracht.

5 *Besluit*

In de brochure Over Leerplanonderzoek behandelt Bijl (1966) als centrale vraag hoe het programma tot stand moet komen voor wat we tegenwoordig de basisvorming noemen. Zijn standpunt is dat de opstelling van een dergelijk programma geen schrijftafelwerk meer kan zijn, ook niet van een commissie van deskundigen, maar dat er onderzoek nodig is. 'Het onderzoek moet dus uitgaan naar wat er gekend en gekund wordt, maar er moet ook – democratisch! – gevraagd worden naar wat men wenselijk acht. Ten slotte beslist echter over wat aan correctie op het aangetroffen beeld wordt aangebracht, het opvoedingsgezag, d.w.z. de organen die met dat gezag zijn bekleed.' Hiermee geeft hij een uitstekende rationale avant la lettre voor het peilingsonderzoek in Nederland zoals dat in het voorgaande is beschreven.

PPON is te beschouwen als een diepte-investering om politiek en maatschappelijk gewenste processen mogelijk te maken. Uiteraard moet peilingsonderzoek op een technisch adequate manier worden uitgevoerd en in dit opzicht zullen ongetwijfeld verbeteringen mogelijk blijken te zijn. Het succes van PPON moet echter vooral blijken uit de manier waarop de gegevens gebruikt worden ten behoeve van de ontwikkeling van het onderwijs.

Noten

1. Inmiddels zijn ook proefonderzoeken in het speciaal en het voortgezet onderwijs uitgevoerd (zie Vergeer, Oostdam & Otter, 1988; Kuhlemeier & Van den Bergh, 1989) en de minister heeft het Cito medio 1989 opdracht verleend de peilingen naar het speciaal onderwijs uit te breiden. De consequenties van dit besluit, dat na de afsluiting van het manuscript bekend werd, blijven in dit artikel nog buiten beschouwing. In principe zal de eerste peilingsronde in het speciaal onderwijs – vooralsnog beperkt tot LOM- en MLK-scholen – een geheel gaan vormen met de tweede ronde in het basisonderwijs.

Literatuur

Bock, R. D., R. Mislevy & C. Woodson, The next stage in educational assessment. *Educational Researcher*, 1982, *11*, no. 3, 411, 16.

Bokhove, J., H. Brink & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (2). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1988a, *6*, nr. 3, 3-22.

Bokhove, J., H. Brink & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (3). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1988b, *6*, nr. 4, 47-57.

Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs. *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1987, *6*, nr. 1, 3-6.

Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (4). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1988c, *7*, nr. 2, 16-34.

Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (5). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1989a, *7*, nr. 3/4, 3-14.

Bokhove, J. & J. Janssen, Peiling van het rekenonderwijs. *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1989b, *8*, 66-70.

Bokhove, J. & J. Janssen, PPON-onderzoek naar oplossingswijzen. *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1989c, *8*, 125-129.

Bokhove, J. & J. Janssen, PPON-onderzoek naar oplossingswijzen (2). *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1989d, *8*, 184-189.

Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (6). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1989e, *8*, nr. 2, 3-32.

Brink, H., J.M. Wijnstra & F.J.G. Janssens, *Inventarisatie van het leer- en vormingsaanbod bij het bewegingsonderwijs in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 1).

Brink, H., J.M. Wijnstra & F.J.G. Janssens, *Inventarisatie van het leer- en vormingsaanbod bij de beeldende vorming in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 2).

Bijl, J., *Over leerplanonderzoek*. Groningen: Wolters, 1966 (Acta Paedagogica Ultrajectina XXIII).

Doornbos, K., *Opstaan tegen het zittenblijven*. 's-Gravenhage: Staatsuitgeverij, 1969.

Eggen, T.J.H.M., Itemcalibratie bij vertakt toetsen. In: W.J. van der Linden & L.J.T. van der Kamp (red.), *Meetmethoden en data-analyse* (pp. 75-88). Amsterdam: Swets & Zeitlinger, 1989.

Eisner, E.W., Instructional and expressive objectives: their formulation and use in curriculum. In: W.J. Popham, E.W. Eisner, H.J. Sullivan & L.L. Tyler, *Instructional objectives* (pp. 1-31). Chicago: Rand McNally, 1969 (AERA Monograph Series on Curriculum Evaluation no. 3).

ICB, *Innovatieplan basisonderwijs*. 's-Gravenhage: Staatsuitgeverij, 1978.

Kuhlemeier, H. & H. van den Bergh, *De proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs*. Arnhem: Instituut voor Toetsontwikkeling, 1989.

Kwaliteit van het Onderwijs. 's-Gravenhage: Ministerie van Onderwijs en Wetenschappen, 1981.

Pandey, T.N., State of the art of large-scale assessment in the United States. In: W.J. van der Linden & J.M. Wijnstra (red.), *Ontwikkelingen in de methodologie van het onderwijsonderzoek* (pp. 5-25). Lisse: Swets & Zeitlinger, 1986.

Richards, C.E., A typology of educational monitoring systems. *Educational Evaluation and Policy Analysis*, 1988, *10*, 106-116.

Treffers, A., Beredeneerde eindtermen rekenwiskunde voor het basisonderwijs. *Willem Bartjens; Tijdschrift voor Reken Wiskundeonderwijs in de basisschool*, 1988, *8*, 9-52.

Verhelst, N.D., Informatiewinst bij vertakt toetsen. In: W.J. van der Linden & L.J.T. van der Kamp (red.), *Meetmethoden en data-analyse* (pp. 89-96). Amsterdam: Swets & Zeitlinger, 1989.

Verhelst, N.D. & T.J.H.M. Eggen, *Psychometrische en statistische aspecten van peilingsonderzoek*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 4).

Vergeer, M.M., R.J. Oostdam & M.E. Otter, *Taalonderwijs op LOM- en MLK-scholen; voorstudie periodieke peiling van het onderwijsniveau*. 's-Gravenhage: Het Instituut voor Onderzoek van het Onderwijs (SVO), 1988 (Selecta-reeks).

Wesdorp, H., *Goed onderwijs, wat is dat?* 's-Gravenhage: Stichting voor Onderzoek van het Onderwijs, 1985 (Selecta-reeks).

Wesdorp, H., H. van den Bergh, D.J. Bos, J.B. Hoeksma, R.J. Oostdam, J. Scheerens & B. Triesscheijn, *De haalbaarheid van periodiek peilingsonderzoek; een voorstudie op het gebied van het taalonderwijs in de lagere school*. Lisse: Swets & Zeitlinger, 1986 (SVO-reeks nr. 85).

WRR, *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij, 1986.

Wijnstra, J.M., Wat leren kinderen op school? *Pedagogische Studiën*, 1982, *59*, 223-232 en 277-287.

Wijnstra, J.M. (red.), *Balans van het rekenonderwijs in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1988.

Wijnstra, J.M. (red.), *Verantwoording van de rekenpeiling medio en einde basisonderwijs 1987*. Arnhem: Instituut voor Toetsontwikkeling, 1990 (PPON-rapport nr. 5).

Zwarts, M.A. & J.M. Wijnstra, *Inventarisatie van het leer- en vormingsaanbod bij het muziekonderwijs in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 3).

Curriculum vitae

J. M. Wijnstra (1946) studeerde na zijn opleiding tot onderwijzer pedagogiek aan de Rijksuniversiteit te Utrecht (specialisatie onderwijskunde) en promoveerde in 1976 op het proefschrift 'Het onderwijs aan van huis uit friestalige kinderen'. Sinds 1978 is hij als medewerker verbonden aan het Instituut voor Toetsontwikkeling (Cito) te Arnhem, tegenwoordig als projectleider van het project Periodieke Peiling van het Onderwijsniveau (PPON).

Adres: Instituut voor Toetsontwikkeling (Cito) Postbus 1034, 6801 MG Arnhem

Manuscript aanvaard 6-12-'89

Summary

Wijnstra, J. M. 'The design of the Dutch national assessment program in education illustrated with examples from the first mathematics assessment.' *Pedagogische Studiën*, 1990, *67*, 139-150.

The operational preparations for the Dutch national assessment program in education started in 1986. The aim of the program is to periodically describe the contents of the curriculum in use and the achievement of students in primary education, so as to provide a basis for a more rational discussion about the contents and the level of education. In this paper the aim and background of the program are explained in relation to recent policy on a national core curriculum. Subsequently, the organization, planning, and procedures of the project are described. The procedures are illustrated with examples from the first mathematics assessment at the end of primary education, conducted in 1987. In this section special attention is paid to the way the findings are reported, in order to foster the public discussion of these issues.

De progressie in de Nederlandse evaluatiepraktijk

W. J. NIJHOF en J. SCHEERENS
Faculteit der Toegepaste Onderwijskunde
Universiteit Twente

In deze slotbijdrage willen we beknopt terugblikken op een aantal saillante aspecten van de beschreven evaluatiestudies, en wel waar het betreft de produktgerichtheid en de organisatorische configuratie van de evaluatiestudies, de designs, en de bruikbaarheid van de resultaten.

Produkt- versus procesevaluatie

In de discussies over de kwaliteit van het onderwijs is de laatste jaren steeds duidelijker het accent komen te liggen op de resultaten van het onderwijssysteem en in het bijzonder op de effecten van voorgenomen of voor te nemen innovaties (Scheerens, 1989). De bijdragen van Peschar, van Van der Werf en Tesser, en van Wijnstra verschillen uiteraard in aanpak, maar beogen in wezen leereffecten op te sporen van interventies in het onderwijssysteem.

Indien we de dimensie produktgerichtheid versus procesgerichtheid als ordeningscriterium hanteren dan vallen de bijdragen gericht op de evaluatie van het INSP duidelijk onder de procesgerichtheid. Dat ligt ook wel voor de hand, omdat het INSP veel meer als een stimulerend ontwikkelingsprogramma is aangediend dan als een programma waarvan de effecten op het niveau van leerresultaten van meet af aan duidelijk waren. Het doel was niet het meten van leerprestaties maar nagaan of de beoogde managementdoelen gehaald waren. Het INSP kent een lange implementatiefase, voor leereffecten in termen van leerprestaties was het nog te vroeg.

De andere evaluatiestudies, PPON, OVB en Basisvorming, hebben een uitgesproken en expliciet doel, nl. vaststellen wat het effect van onderwijs is aan de hand van metingen (peilingen) en wel in vergelijkende zin op het gebied van taal, rekenen, wiskunde, natuurwetenschappen, etc.

Ten behoeve van basisvorming is, door alle controverses, nog geen langdurige vergelijkende studie uitgevoerd. Peschar geeft aan dat ondanks het feit dat basisvorming nog niet is ingevoerd vooraf een zinvolle discussie mogelijk is over de wijze van effectmeting. Het uiteindelijk doel van de door hem gepresenteerde studie is een uitspraak te kunnen doen over de relatie tussen leereffecten, leerling- en schoolkenmerken. Daaruit valt af te leiden of basisvorming ertoe doet of niet. Dat biedt een zakelijke grondslag voor een discussie over basisvorming. De ideologische discussie over gelijkheid en ongelijkheid kan er vermoedelijk mee geneutraliseerd worden.

Wat we kunnen vaststellen is dat de eenzijdige georiënteerdheid op processen en op beschrijvend onderzoek alleen verlaten is en aangevuld wordt met voorstudies, vormen van analytische evaluatie, evalueerbaarheidsstudies, quasi-experimentele designs en assessmentstudies: vormen van wetenschappelijk onderzoek die gebaseerd zijn op doelrationaliteit (Beem & Peschar, 1988; Hofstee, 1989).

De hier gepresenteerde studies geven blijk van die georiënteerdheid, al boeten sommige daarop in door de lastige organisatorische configuraties waarbinnen ze moeten opereren.

De organisatorische configuratie

Evaluaties van innovaties worden in ons land door een groot aantal actoren in een bureaucratische greep gehouden. Ministeries, verzorgingsstructuur (SVO, CITO, SLO en LPC's), (para-)universitaire instituten, en onafhankelijke door de Minister van Onderwijs ingestelde commissies maken evaluatief Nederland soms tot een vermakelijk gezelschap aandeelhouders. Deze configuratie leidt ook tot langdurige vormen van overleg, waarin soms minder de kwaliteit van evaluatiedesigns ter discussie staat dan de vraag wie wat behoort te beheren. Daarnaast speelt politieke bias een rol (Scheerens, 1989).

Hoe complex en omvangrijk het Informatica Stimuleringsproject ook was, opvallend is en was dat voor deze onderneming, een interdepartementale projectgroep kon worden gevormd die leiding kon geven aan het geheel. Dat is een unicum in het overheidsbeleid, ook waar het betreft de relatief gunstige samenwerking en uitvoering in de evaluatie. Dat ligt beduidend moeilijker en gevoeliger bij de andere projecten, waar doorgaans vele 'afstemmingsgesprekken' noodzakelijk zijn. Het is kennelijk (nog) niet mogelijk om één deskundig instituut of persoon te belasten met een grootschalige evaluatie. Tyler leidde de nationale Eight Year Study in de Verenigde Staten vanuit de University of Chicago in de jaren dertig. Fend superviseerde de grote vergelijkende evaluatiestudies van de Gesamtschule in de jaren zestig in West Duitsland. PPON heeft in de loop der jaren een duidelijke plaats toegewezen gekregen bij het CITO.

Basisvorming en OVB zijn treffende voorbeelden van langdurige interacties.

Nieuwe evaluatiefenomenen

Dat die interacties óók vruchten kunnen afwerpen blijkt met name bij die beide projecten. Zo startte in 1988 een procedure om te bepalen hoe een evaluatie van de basisvorming zou kunnen worden bedacht. Daartoe werd een analytische evaluatie opgezet, een gedachten-experiment waarbij op grond van beredeneerde keuzen door deskundigen vanuit een drietal disciplines evaluatievragen en opzetten werden beschreven (Beem, 1988). Het door de overheid aanvaarde plan voor de basisvorming (WRR, 1986) gaf duidelijk aanleiding tot een dergelijke actie. Deze analytische evaluatie klinkt door in de bijdrage van Peschar, met name waar de criteriummaten en het design aan de orde worden gesteld.

Het fenomeen van de voorstudie is niet nieuw. Opvallend is wel dat de gepresenteerde evaluatieprojecten alle voorstudies kenden op grond waarvan de uiteindelijke opzetten zijn aangepast of verbeterd. De INSPstudie (zie Scheerens' bijdrage) laat zien dat er een evalueerbaarheidsstudie heeft plaatsgevonden met als doel na te gaan of het zinvol zou zijn om te evalueren.

Een evaluatie van analytische voorstudies en evalueerbaarheidsstudies zelf is betekenisvol voor de verdere ontwikkeling van de evaluatiemethodologie.

Aantoonbaar is dat de rationaliteit van beslissen t.a.v. objecten van evaluatie en de context waarin dat moet gebeuren, door deze nieuwe fenomenen aanzienlijk is vergroot. Het verdient aanbeveling bij grootschalige studies de tijd te nemen voor deze exercities. Ze vergroten de kans op doelmatige opzetten, uitvoeringsmogelijkheden en – naar verwacht mag worden – vertrouwen bij politici.

De designs en de criteriummaten

Bij een vergelijking van de opzetten van enkele studies valt op dat een model voor schooleffectiviteitsonderzoek heeft gefunctioneerd (Walberg, 1984; Frazer, 1987) dat leerlingkenmerken, achtergrondkenmerken en schoolkenmerken als interactief concept hanteert. De eenduidige lineaire relatie tussen leerlingkenmerk en schoolprestatie als uitgangspunt is daarmee verlaten. Zo blijkt in de studies van Peschar en van Van der Werf en Tesser heel nadrukkelijk gekozen te zijn voor cohortstudies met een vergelijkend karakter en een multi-level analyse.

Wijnstra schetst een steekproefonderzoek (matrix sampling) als voorbereiding op meer diepgaande case studies op scholen om deze te helpen tot zelfevaluatie te komen. Matrix sampling betekent dat niet iedere leerling of klas alle opgaven van een meetinstrument maakt, terwijl onder bepaalde voorwaarden over alle scholen gegeneraliseerd kan worden. De koppeling tussen matrix sampling en terugkoppeling naar scholen waarvoor Wijnstra pleit, staat door deze procedure onder druk. Terugkoppeling naar scholen is immers pas doelmatig wanneer alle gegevens van alle leerlingen van alle af te nemen toetsen van die school in het databestand kunnen worden teruggekoppeld. Die garantie is niet op voorhand aanwezig. Vandaar dat case studies worden uitgevoerd om na te gaan hoe die terugkoppeling wel gerealiseerd kan worden. Een tweede element is dat de peilingen curriculumonafhankelijk zijn en vooral een functie hebben in het licht van te ontwikkelen en dus te toetsen eindtermen. Dat betekent ook dat die terugkoppeling van gegevens nog niet tot consequentie heeft dat het onderwijsleerproces kan worden verbeterd.

De criteria die in de verschillende onderzoeken worden aangelegd om effecten vast te stellen vormen niet uitsluitend leerprestaties. Bij het OVB speelt b.v. het studieadvies in de voorstudie een zware rol. Deze wordt gecorreleerd met de wijze waarop de OVB-wegingsfactor intern in de organisatie is vertaald in personeel, begeleiding e.d. Bij latere evaluaties wordt veel nadrukkelijker op cognitieve leerprestaties gelet.

Bij PPON speelt de prestatie op een meetinstrument een rol. In de onderhavige studie voor rekenen. Een absolute norm is er (nog) niet voor het aanvaardbare rekenpeil. Opvallend is wel dat het vinden van die norm niet eenvoudig lijkt, waar de beoordelaarsgroepen significant van elkaar verschillen qua niveau inschatting. De normatieve schattingen van alle groepen gaan ver uit boven wat de leerlingen feitelijk laten zien. Hier blijkt hoe lastig het is eindtermen eenduidig operationeel op een bepaald niveau te formuleren en als zodanig te laten beoordelen.

Bij het onderzoek naar basisvorming staat de vraag centraal of scholen met basisvorming gekenmerkt worden door een verhoging van de kwaliteit van het onderwijs onder constant houding van leerlingkenmerken. Kwaliteit wordt gemeten in termen van prestatie- en niet-cognitieve toetsen. Aandacht voor omgevingskenmerken bestaat daarin dat de invloed van het curriculum en van de organisatie gemeten wordt. Op deze wijze wordt doelrationeel bezien of een beoogde innovatie werkt. Het regeerakkoord van CDA en PvdA van november 1989 heeft inmiddels een lastig fenomeen uit de basisvorming gehaald: eindtermen op twee niveaus. Dat maakt de evaluatie daarvan en de technische invoering ietwat eenvoudiger, de aanvaardbaarheid voor scholen vermoedelijk niet.

De managementeffecten van het INSP zijn niet onverdeeld ongunstig. Er is een forse stap vooruit gezet waar het betreft Nederland te 'informatiseren'. Of dat alles voor de scholen zelf nu de grote stap voorwaarts is geweest dient op grond van de resultaten genuanceerd te worden bezien. Er zijn te grote verschillen in projecten en programma's om daarover een generaal oordeel te vellen. De trein is in beweging gezet. En daarmee komen we ten slotte bij het laatste element.

Bruikbaarheid en het gebruik van de evaluatiestudies

Doel van dit themanummer was te laten zien dat er enige progressie schuilt in de evaluatiestudies en opzetten. Nieuwe fenomenen als analytische evaluatie, evalueerbaarheidsanalyses, doordachte opzetten in longitudinaal perspectief met multi level analyses betekenen een sprong vooruit vergeleken met de traditioneel beschrijvende processtudies. Het denken in termen van effecten is dominanter en duidelijker geworden. Daarmee zijn de problemen niet opgelost, want de designs zijn complexer geworden en vatbaar voor bedreigingen. Voor de wetenschappelijke theorievorming geldt echter dat sterke designs uiteindelijk meer inzicht zullen geven in de wijze waarop onderwijssystemen functioneren. Daardoor kan beïnvloeding van beleid en van scholen zelf op termijn ook doelmatig zijn. Evaluatie is uiteindelijk bedoeld voor de kwaliteitsverbetering van het onderwijs.

Literatuur

Beem, A.L. (red.), *Weten, Redeneren, Raden: een analytische evaluatie van de basisvorming*. Lisse: Swets & Zeitlinger, 1988.

Fraser, B.J. et al., Syntheses of educational productivity research. *International Journal of Educational Research*, 1987, *11*, 2, 145-252.

Hofstee, W.K.B., Naar een zuiver wetenschappelijke opvatting van evaluatie-onderzoek. In: J. Scheerens (red.), *Evaluatie: om de kwaliteit van het onderwijs*. Lisse: Swets & Zeitlinger, 1989.

Peschar, J.L., OP-STAP. Enkele implicaties van de analytische evaluatie voor het evaluatieplan voor het voortgezet onderwijs eerste fase (basisvorming). In: A.L. Beem (red.), *Weten, Redeneren, Raden: een analytische evaluatie van de basisvorming* (pp. 85-96). Lisse: Swets & Zeitlinger, 1988.

Scheerens, J. (red.), *Evaluatie: om de kwaliteit van het Onderwijs*. Lisse: Swets & Zeitlinger, 1989.

Walberg, H.J. Improving the productivity of America's schools. *Educational Leadership*, 1984, *41*, 208, 19-30.

WRR *Basisvorming in het Onderwijs*. Den Haag: SDU, 1986.

Boekbespreking

J. M. M. van der Sanden (red.), *Huiswerk en huiswerkbeleid* (S.V.O.-reeks 87). Swets & Zeitlinger, Lisse, 1989, 235 pag. + 39 pag. bijlagen, *f*51,50, ISBN 90 265 1022 5.

Huiswerk is niet een thema dat onderwijskundigen en spraakmakende onderwijsgevenden frequent heeft beziggehouden blijkens de betrekkelijk geringe omvang (± 130 publikaties, waarvan ± 70 van Nederlandse oorsprong) van de literatuur waaraan in dit boek is gerefereerd. De publikaties waarvan er een 38-tal achtergrondliteratuur vormen en niet het huiswerk of het beleid daaromtrent zelf betreffen, bestrijken de periode van circa 1960 tot het eind van het onderzoek dat in dit boek wordt beschreven (31 dec. 1987), al zijn er ook nog een achttal uit 1988 en 1989, de verslagperiode, opgenomen.

Ook Nederlandse ouders, in tegenstelling tot die uit de BRD, bemoeien zich weinig met het huiswerk van hun kinderen (p. 15 en p. 46). Waarom dan toch een onderzoek eraan gewijd?

De laatste jaren schijnen regelmatig publikaties te zijn verschenen die melding maken van problemen van leerlingen met het huiswerk, zo frequent kennelijk dat Tweede Kamer-lid Franssen (VVD) vragen heeft gesteld die de gewezen staatssecretaris Ginjaar-Maas ertoe brachten een onderzoek aan te kondigen en op te dragen naar het huiswerkbeleid van scholen voor voortgezet onderwijs.

Dit boek vormt het verslag van dit door de K.U.B. in samenwerking met de R.U.L. (Dr. Neuwahl) verrichte onderzoek, dat bestaat uit twee delen: een literatuurstudie en een empirisch onderzoek.

Het literatuuronderzoek beslaat een negental hoofdstukken betreffende opvattingen over huiswerk, de relatie tussen huiswerk en leerprestaties, de functies van huiswerk, de hulp erbij op school en thuis, de manier waarop leerlingen hun huiswerk aanpakken, aanbevelingen over het maken ervan alsmede over de wijze waarop docenten met huiswerk omgaan en over het huiswerkbeleid van scholen.

Hoewel het opgeven van huiswerk in het algemeen heel gewoon wordt gevonden en huiswerk ook zinvol wordt geacht door ouders, leerlingen en docenten, doet zich toch een aantal problemen voor die o.m. te maken hebben met divergente opvattingen over functies van huiswerk. Deze worden door de onderzoekers met name in het voetspoor van Kamm in een leertheoretisch kader geplaatst: wekken van leerinteresse en leren zelfstandig werken (oefening, toepassing en ontdekkend leren). Merkwaardigerwijs berusten de gegeven aanbevelingen (hoofdstuk 7) nog vooral op vuistregels. Empirisch onderzoek is dus wel geïndiceerd. Te meer omdat van een huiswerkbeleid op scholen in de literatuur in het algemeen geen sprake lijkt te zijn, terwijl zulk een beleid wel zinvol wordt geacht.

Deel II verslaat empirisch onderzoek met een zevental centrale vragen: of scholen een samenhangend huiswerkbeleid voeren, hoe ze daarin eventueel verschillen, of er samenhang is tussen huiswerkmoeilijkheden, studiegewoonten en huiswerkaanpak van leerlingen, of verschillend schoolbeleid op dit terrein samenhangt met verschillen in leerprestaties, huiswerkaanpak, huiswerkmoeilijkheden en studiegewoonten van leerlingen en of er ten aanzien van die vragen verschillen zijn tussen schooltypen. Het onderzoek werd na een voorbereidend interviewonderzoek uitgevoerd met drie vragenlijsten voor directies, studielescoördinatoren (vanwege de eventuele samenhang tussen huiswerk en studielessen) en de sectie coördinatoren van een zaakvak, een taal, een exact vak met het oog op huiswerkbeleid.

Wat de huiswerkaanpak betreft werd de Huiswerkvragenlijst (HAVT), die theoretisch werd gefundeerd in de handelingsregulatietheorie die aan Kuhl werd ontleend, door de K.U.B. geconstrueerd. Huiswerkmoeilijkheden en studiegewoonten werden in kaart gebracht met de reeds eerder geconstrueerde Vragenlijst Huiswerk Voortgezet Onderwijs (VHVO) van de R.U.L. Alle gegevens betreffende de constructie van de huiswerkbeleidvragenlijsten komen ter sprake in hoofdstuk 3 van deel II en die over de huiswerkaanpakvragenlijst in hoofdstuk 7. De procedure van constructie blijkt op zeer verantwoorde wijze te zijn uitgevoerd. In de empirisch zeer interessante hoofdstukken 8 en 9 komen achtereenvolgens de resultaten met de HAVT en met de VHVO en de relaties tussen die lijsten aan de orde, alsmede de samenhangen tussen de variabelen uit de verschillende beleidsvragenlijsten, de antwoordpatronen op die beleidsvariabelen en de verschillen tussen scholen qua grootte, type, signatuur en beleid. De bevindingen zijn het meest interessant – en m.i. ook het best theoretisch te onderbouwen – met betrekking tot de schooltypen. Het lbo (in mindere mate ook het mavo) vertoont het meest een gebrek aan een samenhangend huiswerkbeleid en het minst een integratie van studielessen in het onderwijs. Interessant is ook dat geen relatie is gevonden tussen de vragenlijstscores op de HAVT en de VHVO en de leerresultaten uitgedrukt in schoolcijfers.

Pedagogische Studiën 1990 (67) 154-156

Het samenvattende hoofdstuk 10 bespreekt de beantwoording van alle centrale vragen behalve vraag 7, die kennelijk in hoofdstuk 9 geacht wordt voldoende aan bod te zijn geweest. Hoofdstuk 11 bevat de conclusies en de aanbevelingen.

Het beeld dat van het huiswerkbeleid in het conclusiehoofdstuk wordt geschetst is dat het huiswerk als onderdeel van de schoolpraktijk niet ter discussie staat, dat een samenhangend gericht beleid slechts sporadisch voorkomt, en dat terwijl huiswerk voor leerlingen geenszins een probleemloze aangelegenheid is.

Er bestaat een discrepantie tussen de ernst van de gesignaleerde problemen van planning, aanpak, moeite met huiswerk, vooral leerwerk, ongemotiveerdheid, controle en concentratie van de kant van de leerlingen en de reactie van de school daarop. Dat is zozeer het geval dat de conclusie uitmondt in de constatering 'dat er met betrekking tot huiswerk sprake is van een situatie waarin elk van de betrokkenen primair op zichzelf is aangewezen voor het vinden van oplossingen voor de problemen' (p. 208). Dat vooral de leerlingen daarvan de dupe zijn ligt voor de hand.

De aanbevelingen richten zich op de school, zowel op centraal niveau als op dat van de secties en van de individuele leraren. Zelfonderzoek ten aanzien van het huiswerkbeleid en een andere inrichting van de studielessen (trainingsprogramma's voor leerlingen in zelfstandig leren), ook met het oog op de toekomstige basisvorming, zijn de aanbevolen recepten. Het is merkwaardig dat de ouders er niet in betrokken worden, terwijl op p. 88 in de literatuurstudie wordt geconstateerd dat er onvoldoende communicatie is over huiswerk tussen ouders en de school.

'Huiswerk en huiswerkbeleid' is een goed geschreven en goed leesbaar boek, dat theoretisch en empirisch goed onderbouwd is. Ieder hoofdstuk en vele afzonderlijke paragrafen zijn voorzien van korte, heldere samenvattingen. De hoofdstukken 10 en 11 van het empirische deel, maar ook de hoofdstukken van het literatuuronderzoek zijn zonder veel inspanning door niet-experts in onderzoek (onderwijsgevenden, ouders) te lezen. De andere hoofdstukken over de procedure van het empirisch onderzoek zijn voor wie zich wat meer wil inspannen en de verantwoording van de samenvatting, de conclusies en de aanbevelingen wil nagaan, goed leesbaar, ook als ze geen experts zijn. Daarom is dit boek voor alle geïnteresseerden in huiswerkproblematiek, met uitzondering van de leerlingen waarschijnlijk, een welkome vraagbaak. Van harte aanbevolen.

A. M. P. Knoers

W. Koops & J. J. van der Werff (red.), *Overzicht van de empirische ontwikkelingspsychologie. 1. Grondkenmerken van het vak.* Wolters-Noordhoff, Groningen, 1987, 220 pag., *f*45,00, ISBN 90 01 48896 X.

In 1979 verscheen het *Overzicht van de ontwikkelingspsychologie*, samengesteld door W. Koops en J. J. van der Werff met de bedoeling om de empirische ontwikkelingspsychologie voor een Nederlands publiek op systematische wijze te presenteren. In plaats van een tweede druk is nu een herziene versie verschenen in drie delen: deel 1 behandelt de 'grondkenmerken van het vak', deel 2 gaat over 'de ontwikkeling van functies en cognitie' en deel 3 over 'sociale ontwikkeling en de ontwikkeling van de persoonlijkheid'. In deze bespreking beperk ik mij tot deel 1.

Het boek bevat zes hoofdstukken, vier daarvan (van Koops en Van der Werff, van Baerends, van Bronfenbrenner, en van Van der Werff) waren al in de editie van 1979 te vinden (zij het dat zij voor deze nieuwe editie met wat aanvullingen up to date gemaakt zijn); twee hoofdstukken verschijnen voor het eerst (die van Van Geert en Streng).

Het inleidende hoofdstuk 1, geschreven door Koops en Van der Werff, bevat een schets van de ontwikkelingspsychologie. De geschiedenis van het vak komt beknopt ter sprake; de belangrijkste huidige trends en aandachtsverschuivingen worden genoemd. In vergelijking met de versie uit 1979 gaan de auteurs meer in op de relatie van de ontwikkelingspsychologie met de pedagogiek en de kinderpsychiatrie. Van Geert analyseert, in hoofdstuk 2, de betekenis van het begrip 'ontwikkeling' en de structuur van ontwikkelingspsychologische theorieën.

Hoofdstuk 3 en 4 zijn in het boek opgenomen om de twee polen van het denken over ontwikkeling weer te geven: de biologische invalshoek en het omgevingsdenken. Baerends geeft in een hoofdstuk over de ontwikkelingsbiologie van het gedrag een beeld van het ethologisch onderzoek en gaat in op het begrip organisatiestructuur van gedrag. Bronfenbrenner houdt in een principiële beschouwing een pleidooi voor 'ecologisch' onderzoek waarin de wederzijdse aanpassing van individu en omgeving wordt onderzocht. Ecologisch onderzoek moet zich, in de visie van Bronfenbrenner, niet beperken tot de invloed van de onmiddellijke omgeving, maar dient ook de rol van culturele ideologieën en sociale instituties (het 'macrosysteem') aan de orde te stellen.

In hoofdstukken 5 en 6 komen een tweetal tradities van empirisch onderzoek aan het woord. Het hoofdstuk van Streng behandelt de stand van zaken in de gedragsgenetica en het hoofdstuk van Van der

Werff geeft een overzicht van het onderzoek over de stabiliteit van persoonskenmerken. In dit laatste hoofdstuk wordt ook de belangrijke methodologische kwestie van de voor- en nadelen van longitudinaal en transversaal onderzoek uitvoerig belicht.

Het boek is bedoeld voor lezers die al een basiskennis van de ontwikkelingspsychologie bezitten. Het is niet geschikt als eerste kennismaking met het vak; daarvoor wordt teveel bekend verondersteld over theorieën en methoden. Gevorderde studenten en pedagogen en kinderpsychologen die hun kennis van de ontwikkelingspsychologie willen verdiepen zullen in dit boek veel van hun interesse aantreffen. De hoofdstukken zijn goed toegankelijk, en de schrijvers weten moeilijke problemen over het algemeen zonder jargon te verduidelijken.

Het boek als geheel vind ik minder goed geslaagd. Mijn bezwaren richten zich tegen de keuze van de onderwerpen die ter sprake komen. Naar mijn smaak is de biologische kant van de ontwikkeling (in de hoofdstukken van Baerends en Streng) overtuigender gerepresenteerd dan de invloed van omgeving en cultuur. Het hoofdstuk van Bronfenbrenner kon in 1979 een belangrijke rol vervullen omdat het aangaf hoe omgevingsvariabelen, anders dan het behaviorisme dat deed, geconceptualiseerd kunnen worden. Nu, tien jaar later, maakt het teveel de indruk alleen maar een programma te zijn, dat laat zien hoe het zou moeten. In vergelijking met 1979 is het onderzoek in de traditie van de cultuurhistorische school en het crossculturele onderzoek zo ver voortgeschreden dat een principiële bespreking van de betekenis van dit onderzoek op zijn plaats was geweest. De naam Vygotskij, wiens ideeën over de maatschappelijke kanten van de ontwikkeling van het kind het afgelopen decennium voor velen een bron van inspiratie zijn geweest, valt nergens in het hele boek. Zo lijkt het erop dat de biologische benadering van de ontwikkeling voortdurend vooruitgang boekt, terwijl de omgevingsdenkers niet verder komen dan een programmatisch raamwerk.

Een ander probleem komt voort uit de keuze van de samenstellers vooral het ontwikkelingspsychologische *onderzoek* te laten zien. Zij hebben zich niet ten doel gesteld theoretische benaderingen te beschrijven of te belichten. In een boek dat de grondkenmerken van het vak wil behandelen is dat uitgangspunt een nadeel. Als er één 'grondkenmerk' van de ontwikkelingspsychologie is, dan is dat wel het voortdurende gebrek aan overeenstemming over de fundamentele vraag naar het wezen van de ontwikkeling. Wat zijn de oorzaken van de ontwikkeling en hoe moeten wij die ontwikkeling beschrijven? Is de ontwikkeling een kwestie van natuur of van cultuur, van rijping of van leren? Hoe verloopt de ontwikkeling: in stadia of op andere wijze discontinu, of juist geleidelijk? Die vragen liggen ten grondslag aan de grote controversen in de ontwikkelingspsychologie, bijvoorbeeld tussen het behaviorisme en de cognitieve psychologie, of tussen Piaget en Vygotskij. Deze kwesties zijn niet zomaar in empirisch onderzoek te beslechten. De moderne visies op de wetenschap, van Kuhn, Lakatos en anderen, hebben ons geleerd dat elke theorie niet-toetsbare vooronderstellingen bevat. Deze vooronderstellingen hebben empirische consequenties: zij maken concepten en theoretische ideeën mogelijk met behulp waarvan de werkelijkheid wordt geproblematiseerd. Op die manier wordt het empirisch onderzoek in een bepaalde richting gestuurd. Omdat de samenstellers de nadruk willen leggen op de *empirische* ontwikkelingspsychologie blijven deze aspecten onderbelicht.

Het hoofdstuk van Van Geert komt een eind in de goede richting. Maar hij houdt zich meer bezig met formele verschillen tussen de ontwikkelingstheorieën dan met inhoudelijke meningsverschillen. Door te reflecteren op de theorieën over ontwikkeling weet hij de verschillende theorieën in een logisch schema te zetten. Hij laat bijvoorbeeld zien dat sommige theoretici redeneren vanuit een wenselijke eindtoestand van de ontwikkeling (retrospectieve theorieën) terwijl anderen ervan uitgaan dat de ontwikkeling een open einde heeft (prospectieve theorieën). Dat is wel verhelderend, maar Van Geert gaat niet in op de feitelijke controversen die het gevolg zijn van deze verschillen.

Kortom: het boek bevat een zestal interessante artikelen, maar sommige 'grondkenmerken' van de ontwikkelingspsychologie komen te weinig uit de verf.

E. Elbers

In memoriam Professor Dr. M.J. Langeveld

Op 15 december 1989 is te Naarden professor Martinus Jan Langeveld overleden. In deze aflevering van Pedagogische Studiën blikt professor Van Hulst terug op leven en werk van deze onmiskenbaar grote Nederlandse pedagoog en wetenschapper die ook voor ons tijdschrift grote betekenis heeft gehad.

Langeveld is in 1939, op zijn 34ste reeds, lid geworden van de redactie van Pedagogische Studiën, een lidmaatschap dat hij gedurende 32 jaren heeft vervuld tot 1971, waarna hij lid werd van de redactieraad. Tot begin 1986, hij was toen inmiddels 80 jaar, is hij lid gebleven van de redactieraad.

Langeveld is zeer produktief geweest. In dit tijdschrift schreef hij 54 artikelen en nog veel meer boekbesprekingen. Hij las veel, schreef daarover, had een pregnante opvatting en wist deze buitengewoon eloquent te vertolken, als het moest in andere talen. Met verve heeft hij vanuit zijn pedagogische theorie-opvatting een breedheid aan thema's aangepakt, bewerkt, onderzocht en verdedigd. Van Hulst geeft een karakteristieke schets.

De laatste schriftelijke bijdrage van Langeveld aan Pedagogische Studiën is van april 1977 (jaargang 54, nr 4, pp. 123-129), een polemisch stuk dat de Duitse pedagoog Brezinka treft. Bij herlezing valt op hoezeer Langeveld zijn wetenschapstheoretische en fenomenologische denkwijze en zijn pedagogische opvattingen in de strijd brengt en Brezinka als een exponent van wereldvervreemding afdoet, omdat deze de 'empirie' niet verstaat. Of in de woorden van Langeveld zelf "(...) Het is vreemd aan iedere onmiddellijk pedagogische relatie en situatie. Zijn methodologische grondhouding maakt empirische onderzoeken niet reëel mogelijk door de afwezigheid van de grondveronderstellingen van een op handelen en dus situationeel gerichte wetenschap" (o.c. p. 123). Met dit stuk lijkt Langeveld zijn wetenschappelijke boodschap nog een keer te willen uitdragen. Daarna werd het stil.

De redactie herdenkt in hem een erudiet en begaafd wetenschapper die zijn grote talenten dienstbaar heeft gemaakt aan dit tijdschrift en aan de pedagogiek.

Namens de redactie

W. J. Nijhof

Martinus J. Langeveld (1905-1989)

Korte introductie

Beantwoording van het verzoek van de redactie van Pedagogische Studiën om een herdenkingsartikel over Langeveld te schrijven, stelt mij enerzijds voor een gemakkelijke, anderzijds voor een heel gecompliceerde taak. Gemakkelijk, tenslotte heb ik Langeveld een mensenleven lang gekend, we hebben elkaar regelmatig ontmoet en over belangrijke kwesties correspondeerden wij; van zijn publikaties is mij niet veel ontgaan. Gecompliceerd, want juist de kennis van zijn persoon en werk doet mij aarzelen om 'in kort bestek' een schets te geven van hem, die als geen van zijn collega's in de wetenschappelijke pedagogiek op vele plaatsen in de wereld bekendheid genoot.

Reeds als student stak hij ver boven het gemiddelde uit. Hoewel hij Nederlands studeerde, was het Kohnstamm die als eerste zijn kwaliteiten ontdekte. De oplettende lezer van Kohnstamms "Persoonlijkheid in Wording" (1928) zal bemerken, dat de auteur in het voorbericht zijn dank betuigt 'aan de heer M. J. Langeveld jr., die de moeilijke en tijdrovende samenstelling van het register op zich heeft willen nemen'. En het is wederom Kohnstamm, die twee jaar later in het voorbericht van "Schepper en Schepping, de Heilige", opnieuw zijn dank uitspreekt 'aan de heer M. J. Langeveld, Litt.-Cand.'.

De vraag, of deze informatie wellicht enige signaalwaarde heeft, wil ik zonder reserve positief beantwoorden. In de eerste plaats blijkt hieruit, dat de neerlandicus Langeveld reeds vroeg affiniteit moet hebben gehad met de in opkomst zijnde pedagogiek; vervolgens blijkt uit de samenstelling van de registers van Kohnstamms publikaties, dat de wetenschappelijke interesse van Langeveld een wijde horizon had. In 1930 publiceerde Langeveld in het gezaghebbende tijdschrift De Nieuwe Taalgids (jg. 24) een artikel over de z.g. tussenwerpsels. Zijn dissertatie "Taal en denken" (1934) staat, zoals de titel reeds aangeeft, op het grensgebied van twee wetenschappen: linguïstiek en psychologie; het jaar daarvoor had hij "Op weg naar wijsgerig denken" (1933) gepubliceerd; in de wereld der filosofen werd het als een scherpzinnige studie geprezen. Intussen had hij zijn kennis ook buiten onze landsgrenzen vermeerderd: hij studeerde in Hamburg bij William Stern en in Leipzig bij Th. Litt.

Dat hij een reeks van functies bekleedde is eigenlijk een van zelfsprekende zaak: leraar Nederlands aan het Baarns Lyceum (1931-'35); docent in de kinderpsychologie aan het Nutsseminarium (1935); privaat docent in de pedagogiek der rijpere jeugd aan de Universiteit van Amsterdam (1937), buitengewoon hoogleraar te Utrecht (1939).

Hiermede kan ik de Introductie afsluiten, want in die jaren volgen zijn pedagogische publikaties elkaar in snel tempo op; in totaal heeft hij er meer dan 500 op zijn naam staan.

Kohnstamm en Langeveld

Met enkele woorden sprak ik reeds over de relatie tussen de student Langeveld en de hoogleraar Kohnstamm. Dit kan aanleiding geven tot een misverstand. Men zou er namelijk uit kunnen afleiden, dat Langevelds pedagogiek een voortzetting, een uitbouw en een verdieping zou zijn van wat Kohnstamm als pionier was begonnen; maar dit is niet het geval. Het misverstand kan voortkomen uit de grote waardering, die Langeveld voor Kohnstamm had: hij noemde hem bij voorkeur 'Nederlands grootste pedagoog'.

Ik heb het vermoeden, dat Langeveld reeds als jong student wellicht nog enigermate intuïtief gevoeld moet hebben, dat er ook nog een andere mogelijkheid was om een wetenschappelijke pedagogiek te creëren, dan in *Persoonlijkheid in Wording* het geval was; het is namelijk moeilijk aan te nemen, dat Langeveld, toen hij in 1928 het register van dit werk samenstelde, niet tevens de gehele kopij inhoudelijk op zich heeft laten inwerken en daar al een kritisch oordeel over had.

Er zijn nog andere indicaties die deze veronderstelling grond geven. Langeveld, die toch een uiterst vruchtbaar publicist is, heeft nooit een systematische analyse gegeven van Kohnstamms pedagogiek. Bij de dood van zijn leermeester schrijft hij in Wending (7e jg., 1952) een uitstekend artikel over diens wijsbegeerte, maar niet over diens opvoed-

Pedagogische Studiën 1990 (67) 158-162

kunde. Misschien wijst de ondertitel van Kohnstamms *Persoonlijkheid in Wording*, namelijk *Schets ener christelijke opvoedkunde* ons hier een spoor. Langeveld, die reeds vroeg een aanhanger werd van de wetenschappelijke werkmethode der fenomenologie, wil onder geen beding een pedagogiek die op voorhand gestempeld is door theologische, filosofische of andere wereldbeschouwelijke adjectieven; hij acht dit een ongeoorloofde versmalling van wetenschappelijk denken en werken. Dus geen katholieke, calvinistische of mohammedaanse pedagogiek, maar een fenomenologische, die mogelijkheden biedt aan wetenschappers van zeer uiteenlopende levensovertuigingen, om, uitgaande van de oerfenomenen der opvoeding, samen te werken aan een theoretische pedagogiek. Dan is er in het praktisch opvoedkundig handelen nog ruimte genoeg om eigen wegen te gaan.

Dat de fenomenologisch georiënteerde pedagogiek weerstanden opriep, behoeft ons niet te verwonderen. Het zou ons eerder moeten verbazen, indien dit niet het geval was geweest. De Nederlandse pedagogische traditie werd volstrekt beheerst door het samenvallen van pedagogiek en levensbeschouwing. En aangezien de levensbeschouwing nauw gerelateerd is aan filosofie en theologie, kreeg de gehele pedagogiek daardoor niet slechts een sterk filosofisch en theologisch karakter, maar haar uitgangspunt was menigmaal theologisch. Onder erkenning dat er op zichzelf zeker ruimte is voor een filosofie der opvoeding zowel als voor een theologie der opvoeding, weigert Langeveld principieel deze beide synoniem te achten met theoretische pedagogiek. Dit alles impliceert, dat van Langeveld niet was te verwachten, dat ook van zijn hand een vernieuwde 'schets van een christelijke opvoedkunde' zou verschijnen.

De genese van de fenomenologische methode

Reeds uit de eerste druk van diens *Beknopte theoretische pedagogiek* (1944) blijkt, dat zijn invalshoek een geheel andere is: de pedagogische vragen spruiten niet voort uit de wijsgerige, maar een zuiver pedagogisch betoog roept wijsgerige vraagstellingen op. Hier wordt dus duidelijk het primaat opgeëist voor de pedagogische problematiek, daarbij de wijsgerige en levensbeschouwelijke implicaties naar een veel later stadium verschuivende. Wie dus geen behoefte heeft om in zijn opvoedkundig denken zijn levensbeschouwing op voorhand te expliciteren en daardoor met een betrekkelijk lange traditie breekt, riskeert het tegenspel. Dat is op zichzelf een verantwoorde wetenschappelijke activiteit; maar de kritiek overschrijdt hier zeer licht de grenzen van het wetenschappelijke om te treden in het domein van het persoonlijke. Wie niet meer de eigen geloofsovertuiging tot basis en uitgangspunt neemt voor wetenschappelijk pedagogisch denken, krijgt niet alleen kritiek op zijn methode van werken, maar ook op zijn geloofsovertuiging. Een dergelijke aanval doet pijn, schaadt – in dit geval zelfs beslissend – de intermenselijke relaties en maakt in feite de voortzetting van de wetenschappelijke discussies onmogelijk. Wij doelen hier op de polemiek die Waterink heeft gevoerd met Langeveld. Misschien is het woord polemiek hier nauwelijks juist, want Langeveld heeft eigenlijk nooit gereageerd op Waterinks offensief, afgezien van een enkele aantekening in de *Beknopte theoretische pedagogiek*. Wie hieruit meent de conclusie te mogen trekken, dat het levensbeschouwelijke bij Langeveld er eigenlijk niet zoveel toe doet, heeft zijn werken slecht gelezen.

Intussen viel duidelijk te constateren, dat de fenomenologische pedagogiek na 1945 zowel in Nederland als daarbuiten terreinwinst boekte. In een aantal dissertaties bleek de fenomenologie als werkmethode goed bruikbaar. Voor mij is het proefschrift van E. Vermeer (Spel en spelpedagogische problemen) nog altijd een schoolvoorbeeld van de toepassing van deze methode. Maar ook Perquin en Strasser noemen zich beiden bij voorkeur fenomenoloog, al vertegenwoordigen zij elk weer een aparte richting.

In het Duitstalige gebied bleken de wetenschappers eveneens erg gevoelig te zijn voor hetgeen Langeveld voorstond; daar sloot de fenomenologie ook meer aan bij de traditie die via Dilthey en Husserl naar Theodor Litt loopt. Vertalingen (en herdrukken) van Langeveld en studies over hem zijn evenzovele bewijzen voor zijn invloed.

Het Angelsaktische taalgebied toonde veel meer reserves. De oorzaak ligt hier waarschijnlijk eveneens in een traditie. De talrijke 'philosophies of education' hebben stuk voor stuk weinig affiniteit met de Europees-continentale fenomenologie. Ze zijn pragmatisch, instrumentalistisch en veelal gedrenkt in

de wijsbegeerte van Dewey. Daar komt nog bij dat de vertaling van 'theoretische pedagogiek' door 'philosophy of education' stellig niet adequaat is. Juist door haar pragmatische ondergrond richt zich de philosophy of education veel meer op de vragen van school, onderwijs, didatiek en methodiek dan wat in Nederland onder opvoedkundige probleemstelling wordt verstaan. Het kind is in de Anglo-Amerikaanse opvoedingsfilosofie in de eerste plaats een schoolkind en nauwelijks gezinskind.

Langevelds begrippenmateriaal

Een analyse van Langevelds pedagogiek leert ons *tenminste* het volgende: Hij heeft, eigenlijk als een der eersten in Nederland, zich scherp rekenschap gegeven van het gehanteerde begrippenmateriaal. Het resultaat daarvan is, dat er geen 'zwevende betekenissen' ontstaan waardoor allerlei misverstanden mogelijk zijn. Wie het over opvoeding heeft, moet beginnen te zeggen, wat hij er precies onder wenst te verstaan. Opvoeden is 'het stellen van handelingen nadrukkelijk ter bereiking van een opvoedingsdoel'; dat doel kan niet anders zijn dan 'de inhoudelijke bepaaldheid van de *levenstaak* van de zelfstandige mens'. *Onverenigbaar hiermee is elke poging tot vervalsing* van *het* mens-beeld. Tevergeefs zal men bij Langeveld de uitspraak zoeken, dat kinderen elkaar opvoeden, daar de opvoedingsgezagsverhouding tussen kinderen niet bestaat. Vandaar ook de scherpe scheiding tussen opvoeding en omgang. Maar de omgang kan elk ogenblik omslaan in opvoeden en is dus een 'pedagogisch gepreformeerd veld'. Reeds het bovenstaande impliceert dat Langeveld zich dus onmiddellijk rekenschap geeft van de begrippen: gezag, gezagsuitoefening, gehoorzaamheid, protest, volwassenheid enz. Zo wordt bijvoorbeeld het gezag gezien als zedelijke verantwoordelijkheid en aansprakelijkheid ten bate van anderen. Zeer verhelderend, vooral als noodzakelijk complement van het fenomenologisch uitgangspunt, is Langevelds beschouwing over de opvoedingsdoelen. Hier krijgen de waarden en de waarderangorde die men in het leven kent, het volle pond. De pedagogiek mist hier beslissingsmacht; deze is voorbehouden aan de wijsgerige mensbeschouwing. De evidente grondstelling van de opvoeding is echter, dat zij geen doel in zichzelf heeft, maar dat het gaat om de volwassenheid. En in twee militante volzinnen stelt Langeveld de grens tussen pedagogiek en axiologie (waardeleer): *De pedagogiek bepaalt daarom in eigen huis, welke levenswaarden of vermeende levenswaarden pedagogisch aanvaardbaar zijn. Pedagogisch aanvaardbaar betekent hier dan: verenigbaar met de waarden die HET kinderleven of DIT kinderleven dienen te bepalen.*

Van deze these uit zal de pedagogiek dus het *menswaardige* completeren met het *kinderwaardige*. Langeveld ziet de opvoedingsdoelen (meervoud!) als een complex gegeven, zich niet lenend voor één algemene formule, al wil hij, in overeenstemming met Guardini, wel van een 'Sinnziel' een 'totaal doel spreken': het is datgene wat de opvoeder, onder alle omstandigheden uiteindelijk in en met het kind hoopt te bereiken.

Vijftien jaar later

Vijftien jaar na de eerste druk van de *Beknopte theoretische pedagogiek* heeft Langeveld zich in zijn *Capita uit de algemene methodologie der paedologie* (1959) nog eens duidelijk rekenschap gegeven van zijn opvatting over de fenomenologische methode. Juist omdat deze uiteenzetting 'zeer beknopt' is, heeft zij aan exactheid gewonnen. Vooral zijn verhouding tot de vader der fenomenologie, Edmund Husserl (1859-1938), heeft Langeveld scherp aangegeven. Het komt in hoofdzaak hierop neer. Langeveld baseert zich niet op Husserls filosofie, maar uitsluitend, en dan nog met enkele belangrijke restricties, op diens methode. Als Husserl bijvoorbeeld beweert, dat hij zonder vooronderstellingen werkt, dan zal Langeveld dit bestrijden. Husserl zoekt nl. de werkelijkheid; hiertoe behoort ook de kennis, en dus is het mogelijk, dat hij veronderstellingen omtrent de werkelijkheid gemaakt heeft, die ook de kennis raken. Hieruit concludeert Langeveld o.i. terecht, dat er dus ook de vooronderstelling is, dat het mogelijk is, dat de kennis zich in de verhouding mens en wereld met een 'gegeven' bezig houdt; daardoor is het ook slechts mogelijk, om van fenomeen te spreken, dat in de reductie, de reiniging, van alle historische en metafysische toevalsgegevens wordt ontdaan; er ontstaat dan een nieuwe kenhouding, 'de epoche'. Wie zich inwerkt in Husserls denken, zal doorlopend stuiten op

een aantal basale begrippen, die significant zijn voor diens fenomenologie. De moeilijkheid daarbij is (maar dit geldt absoluut niet alleen voor Husserl, maar in feite voor elke filosoof) dat Husserl zich het recht toeëigent om deze begrippen met een eigen betekeniswaarde te vullen. Dat geldt niet slechts voor de reeds genoemde begrippen als èpoche en reductie, maar ook voor eidos (idee), Wesensschau, intuïtie, intentionaliteit, Einklammerung, noësis, noëmata, cogito, cogitatum, (dit is het denken als activiteit en de inhoud van het denken, de 'gedachtheden') en eveneens voor de schijnbaar eenvoudige begrippen als fenomeen, object, werkelijkheid, waarheid enz. Het lijkt ons derhalve volstrekt onmogelijk het denken van Husserl te begrijpen, nog afgezien van erover te oordelen en het eventueel te veroordelen, indien men niet diep in diens taalarsenaal is doorgedrongen. Vatten wij de draad in het betoog van Langeveld weer op, dan spreekt hij zich uit over de voor de fenomenologische methode zo belangrijke intuïtie, waaronder hij verstaat 'het kennen van strikt *algemene* aard uit onmiddellijke beleving van het object'. Wie wijsgerige interesse heeft, zou eens de fundamentele verschillen na moeten gaan tussen Descartes, Bergson en Husserl, bij hetzelfde woordgebruik, nl. de intuïtie. Conclusie van Langeveld: de intuïtie is een fundamenteel instrument, dat voortdurend controle eist en dat voortdurend corrigerend moet kunnen werken op z.g. 'exacte' methodes.

De centrale plaats van de antropologie

Wij moeten nu komen tot een aantal antropologische noties in Langevelds pedagogiek. Met opzet spreken wij hier niet over 'de' antropologie van Langeveld. Dit zou immers veronderstellen een uitgewerkte leer over de mens die sui generis wijsgerig is. Een dergelijke leer heeft Langeveld niet gegeven. Wel gaf hij ons een aantal noties omtrent het mensbeeld, die een zeker burgerrecht hebben verkregen, temeer omdat hij wellicht de eerste is geweest, die deze noties onmiddellijk functioneel heeft gemaakt in de antropologie van het kind. Hij stuit daarbij op enkele fundamentele kenmerken van het mensbeeld, die zo absoluut overtuigend voor hem zijn, dat zij het uitgangspunt vormen van elk essentieel pedagogisch denken. Het valt ons daarbij op, dat Langeveld juist ten aanzien van dit uitgangspunt het begrip 'evident' gebruikt, dat we stellig in strikt fenomenologische zin moeten duiden. Wellicht zou Strasser hier zelfs van oerevidentie spreken: alle pogingen om dit antropologisch uitgangspunt te ontkennen, zullen impliciet leiden tot de bevestiging daarvan! Het axioma van waaruit Langeveld pedagogiek constitueert, is het volgende: *Dat de mens een wezen is dat opvoedt, opgevoed wordt en op opvoeding is aangewezen, is zelf een van de fundamenteelste kenmerken van het mensbeeld.* Ondanks de evidentie die hieraan toegekend wordt, zal men nochtans de voorzichtige vraag moeten stellen, of het zonder enige reserve is toegestaan het eigen mensbeeld zo absoluut tot 'het' mensenbeeld te proclameren. Het antwoord op deze vraag is namelijk niet van pedagogische, maar van filosofische aard. Het zou namelijk kunnen zijn, dat enkele existentialistisch georiënteerde denkers een mensbeeld concipiëren, dat in een bepaalde spanningsrelatie staat tot dat van Langeveld. Met de begrippen die een zeker burgerrecht in het pedagogisch denken verkregen, bedoelen wij de volgende: Langeveld acht de opvatting volstrekt constitutief dat de mens een *animal educandum* is; het dier dat opgevoed moet worden. Waarom? Omdat het een *animal educabile* is: een op opvoeding aangelegd dier. Nu is dit animal educandum door zijn geestelijkheid een volstrekt uniek wezen, vandaar dat Langeveld zelf zegt, dat we de animal educandum beter kunnen karakteriseren als *homo educandus*, waarmee in feite een tautologie is uitgesproken; zo is het voor hem dus duidelijk, dat de kenschetsing van de mens als homo educandus, educans en educabilis zijn sociaal karakter impliceert. De grond van deze opvattingen ziet Langeveld in een specifieke dubbelheid in de mensbeschouwing: de mens als mens staat alleen als *deze* mens voor zijn God. Het feit, dat God op deze aarde als kind is begonnen, is een fundamentele component in zijn antropologie van het kind; de opdracht die de mens in deze wereld heeft te vervullen, wijst naar dit feit terug.

Waardering en kritiek

Aan waardering heeft het Langeveld stellig niet ontbroken: talrijke herdrukken van zijn werken, vooral van zijn 'Beknopte'; vertalingen van zijn belangrijkste studies; uitnodigin-

gen voor gastcolleges zelfs in Amerika en Japan; twee eredoctoraten; reeksen van ouders die bij hem hulp zochten voor hun kinderen: Prof. Van der Geld heeft in een boeiend artikel in De Telegraaf (13-01-90) hiervan treffende voorbeelden gegeven; volle collegezalen; tientallen promovendi enz. Eigenlijk is zijn ster stijgende geweest tot hij op 65-jarige leeftijd met emeritaat ging. Dat was juist in de jaren, dat de universiteiten de révolte van de studenten meemaakten met als doel een volledige interne democratisering. Ieder weet, dat zich toen duidelijke marxistische tendenzen openbaarden; de studenten verweten de pedagogiek van Langeveld een tekort aan emancipatorische gerichtheid, gebrek aan maatschappijkritiek, te gering politiek voltage, een te dominante houding van de opvoeder, enz. Ieder weet ook, dat deze révolte in feite is doodgebloed en dat men momenteel het marxisme aan de straatstenen nog niet kwijt kan. Maar in die jaren hadden veel docenten in de sociale wetenschappen het heel erg moeilijk. Doch ook van andere zijde kwam de kritiek. De fenomenologische methode, waarvan Langeveld verwacht had, dat ze onverwoestbaar zou zijn, kwam onder zware druk te staan; de vanuit een wijsgerige optiek gestelde vraag, of het ten aanzien van de vader der fenomenologie, Edmund Husserl, wel verantwoord was, om diens filosofie volstrekt te verwerpen, maar de gehanteerde methode te huldigen won veld; reeds in 1962 had Strasser in zijn Opvoedingswetenschap en Opvoedingswijsheid gezegd, dat het standpuntloze standpunt als basis voor theoretische pedagogiek onhoudbaar was en dat daarin de impressionistische fenomenologie plaats moest maken voor de dialogische. Langevelds beschouwingen over de pedagogische antropologie van het kind werden zeker niet meer algemeen aanvaard, vooral niet door die wetenschappers die zich nog enigszins verwant voelden aan het logisch-positivisme (Wiener Kreis, Ernst Mach): zij weigerden de volgens hen volstrekt onverifieerbare uitspraken over 'het wezen van het kind en over het mensbeeld' te onderschrijven. En natuurlijk kwam daarbij, dat de jongere hoogleraren hun eigen boeken schreven en deze als 'verplichte literatuur' invoerden. Langeveld was duidelijk op de terugtocht. Spiecker gaf in zijn dissertatie over de "Idealen en idolen van de opvoedingswetenschap" (1974) slechts marginale aandacht aan Langeveld, hij schrijft een apart hoofdstuk over het afscheid van de pedagogische antropologie en noemt uitspraken over *de* mens en *het* kind inhoudloos (p. 94). En J.D. Imelman acht in zijn dissertatie "Plaats en inhoud van een personale pedagogiek" (1974) Kohnstamm en Langeveld wel 'van buitengewoon belang', maar...door hun uitgesproken protestantse signatuur pinnen zij het personale in de mens onmiddellijk vast aan 'het ethisch-normatief gewetensvolle gedrag'.

In het Pedagogisch Tijdschrift/Forum voor opvoedkunde, 8e jg., januari 1983, benadrukte ik opnieuw de belangrijke plaats van Langeveld in de opvoedingswetenschap. Uitgaande van een paradigmatische constructie, poneerde ik de stelling, dat enerzijds elk wetenschappelijk paradigma op den duur op zijn beurt weer plaats moet maken voor een nieuw, terwijl anderzijds de scheppers van dat nieuwe paradigma slechts tot hun schade de catharsis kunnen negeren, die het fenomenologisch denken in de pedagogiek teweeg heeft gebracht.

Ik kreeg een enthousiaste brief van Langeveld, waaruit bleek, hoe goed hem deze waardering had gedaan; hij verstrekte mij nog een aantal gegevens, over de invloed van zijn fenomenologisch paradigma in Japan.

Langevelds laatste woorden

Die woorden zijn te vinden in een interview, dat hij ter gelegenheid van zijn 80e verjaardag toestond aan Jacobs, Miedema en Beugelsdijk (Ped. Tijdschrift 10, 1985). Ik geef slechts een kort citaat: 'Een theorie moet geen theorie zijn, die theoretisch is. Het moet een theorie zijn, die antropologisch de ontmoeting van mens tot mens in de opvoedingssituatie gaat analyseren én vooral ook realiseren. Wanneer de onderzoeker zich terugtrekt in het relatief autonome bastion van de wetenschap, vind ik, dat hij of zij in wezen de zaak in de steek laat'.

Na dit interview zweeg Langeveld. De dood van zijn begaafde vrouw, die bij Albert Verwey gepromoveerd was, van zijn dochter en van zoveel vrienden hadden hem mentaal beschadigd. 'In de vereenzaming mis je de ontmoeting', verzuchtte hij. Maar de ontmoeting met zijn geschriften zal nog jarenlang een bron van inspiratie zijn voor ieder die de opvoedingswetenschap ter harte gaat.

J. W. van Hulst

De ontwikkeling van de Computergestuurde Toets voor Automatisering van het Lezen*

A. VAN DER LEIJ, H. SMEETS en V.H.P. VAN DAAL
Sectie Speciale Pedagogiek,
Vrije Universiteit, Amsterdam

Samenvatting

In deze bijdrage aan de thema-serie 'Instructie van leerlingen met ernstige leerproblemen' wordt de ontwikkeling van de Computergestuurde Toets voor Automatisering van het Lezen (COTAL) beschreven als voorbeeld van onderzoek naar de diagnostiek van kinderen met ernstige leerproblemen volgens een taak- en procesgerichte benadering.

Er is een experimentele versie van een computergestuurde toets ontwikkeld die zeven subtoetsen bevat. Deze subtoetsen weerspiegelen de manieren waarop woorden als stimulus kunnen worden aangeboden (auditief; visueel) en diverse wijzen van antwoord geven (beoordelen op gelijkenis; hardop benoemen; natypen). De moeilijkheidsgraad wordt opgevoerd door verkorting van de visuele stimulus dan wel verlenging van de tijd na de auditieve stimulus. Er worden woorden van zes verschillende orthografische structuren gebruikt.

In de gerapporteerde studie worden de toetsresultaten van 37 leerlingen met ernstige leesproblemen en 37 eerste klassers van overeenkomstig leesniveau vergeleken. De nadruk van deze vergelijking ligt op indicaties voor een verschillend verloop van het proces van leren lezen, toegespitst op directe woordherkenning, een van de pijlers onder het automatiseringsproces.

De resultaten met het instrument zijn bemoedigend te noemen. Verschillen tussen zeer zwakke en normale lezers, o.a. gerelateerd aan het gebruik van de articulatorische code, blijken aantoonbaar. Implicaties voor verder onderzoek worden besproken.

* Dit onderzoek is mede mogelijk gemaakt door een subsidie van het Ministerie van Onderwijs en Wetenschappen o.g.v. het Informatica Stimuleringsplan (INSP). De acteurs danken dr. E. J. Kappers voor het kritisch becommentariëren van een eerdere versie van dit artikel.

1 *Inleiding*

Met dit artikel wordt de themaserie over instructie van leerlingen met ernstige leerproblemen afgesloten. Het past in de ontwikkelingslijn van het differentieel-diagnostisch onderzoek naar de kenmerken van het leerproces van leerlingen met ernstige leerproblemen (Van der Leij & Kappers, 1989). De bedoeling van het project is een instrument te construeren dat voor dergelijke leerlingen: a) de *instructiegevoeligheid* aangeeft: de mate waarin instructie effectief is (geweest), kwantitatief uitgedrukt in beheersingsniveaus van criteriumtoetsen, b) de *instructiebehoeften* indiceert, af te leiden uit de onderlinge verhouding van beheersingsniveaus (het 'leesprofiel') (o.a. de behoefte aan gebruik van verschillende modaliteiten van aanbieding (visueel/auditief) en van diverse mogelijkheden tot antwoord geven: hardop benoemen, naschrijven, beoordelen op gelijkenis). Deze behoeften vormen de grondslag voor differentiatie van de instructie. Het instrument kan tevens gebruikt worden voor evaluatie van de effectiviteit van de instructie.

De belangrijkste aanleiding voor het onderzoek is geweest dat instrumenten die voor instructiegericht diagnostisch gebruik geschikt zijn, in het Nederlands taalgebied niet voor handen zijn, althans niet voor het stadium in het leesproces na het prille begin (i.c. het stadium van de automatisering). Kinderen met ernstige leesproblemen kunnen juist in dit stadium stagnaties vertonen in het leerproces (verg. o.a. Van der Leij, 1983; Van Aarle, 1988). Wat er aan instrumenten is, leent zich slechts voor algemene niveaubepaling (bijv. de EMT van Brus & Voeten, 1973) of is toegespitst op voorbereidend en het begin van het

aanvankelijk lezen (Sas & Wieringa, 1982; Struiksma, Van der Leij & Vieyra, 1987). Met de ontwikkeling van het instrument wordt getracht om het instrumentarium aan te vullen.

Dit artikel beschrijft het eerste gedeelte van het project, toegespitst op de vraag die ons primair relevant leek: in hoeverre indiceert het instrument verschillen in het ontwikkelingsverloop m.b.t. automatisering tussen leerlingen met ernstige leesproblemen en kinderen bij wie het leesproces zich op normale wijze ontwikkelt? Die verschillen vormen immers de grondslag voor realisering van het uiteindelijke doel: diagnostische informatie opleveren die aanpassing van de instructie aan de eigenaardigheden van kinderen met ernstige leesproblemen mogelijk maakt.

2 *Achterliggende theorie*

De resultaten van wetenschappelijk onderzoek (verg. o.a. Perfetti, 1985) indiceren dat er bij het leren lezen twee verwerkingswijzen van woorden zijn. In het beginstadium leert de leerling om woorden *om te zetten in de klankvorm* (ook wel fonologisch decoderen genoemd) en vervolgens verder te verwerken (koppelen aan de betekenis en/of het hardop uitspreken van de woorden). Daarnaast ontwikkelt zich in de loop van het leerproces een tweede verwerkingswijze. Steeds meer woorden worden via *directe woordherkenning*, dat wil zeggen zonder voorafgaande omzetting in de klankvorm, gekoppeld aan de betekenis op grond van ervaring in het lezen van die woorden. Beide processen worden in de loop van de eerste onderwijsjaren geautomatiseerd. Dat betekent dat de ervaren lezer in principe alle woorden snel en nauwkeurig kan verwerken. Bij bekende woorden geschiedt dat door directe herkenning, bij onbekende woorden door omzetting in de klankvorm, wat, vergeleken met bekende woorden, slechts een fractie van een seconde minder snel verloopt (Seymour, 1986). Kenmerkend voor dit automatische verloop is dat de processen geen of nauwelijks aanslag meer doen op de aandachtscapaciteit, die derhalve beschikbaar is voor andere activiteiten waarvan begrijpend lezen uiteraard de belangrijkste is.

Kinderen met ernstige leesproblemen blijken geen moeite te hebben met het waarnemen van visuele stimuli op zichzelf, zoals geschiedt bij het beoordelen van visuele patronen (figuren; letters) op gelijkenis. Tekorten treden echter aan het licht wanneer visuele stimuli door middel van omzetting in een of andere klankcode verwerkt moeten worden, het eerste van de voornoemde twee processen (Bouma & Legein, 1980). Reitsma (1983) heeft aannemelijk gemaakt dat (zeer) zwakke lezers daarnaast grote moeite hebben met het zich eigen maken van 'woordspecifieke' kennis – de kennis van de unieke combinatie van lettertekens van een woord – die woordherkenning zonder letter-voor-letter omzetting in de klankvorm mogelijk maakt. (Normale lezers beginnen daarentegen al vroeg in het proces van leren lezen te beschikken over deze tweede verwerkingswijze.) Het gevolg is dat het automatiseren zowel van de omzetting in de klankvorm (het decoderen) als van de directe woordherkenning problematisch verloopt (Spear & Sternberg, 1987).

3 *Overwegingen bij de ontwikkeling van het instrument*

In algemene zin is voor de ontwikkeling van het instrument het idee als uitgangspunt genomen dat leerlingen met ernstige leesproblemen vooral stagnaties vertonen met het leren lezen van (losse)*woorden*. Stagnaties bij het lezen van zinnen en het begrijpend lezen zijn daarvan een afgeleide (Perfetti, 1985). Onze interesse ging primair uit naar het lezen van woorden die door veelvuldig gebruik in de kinderspreek- en schrijftaal qua betekenis in aanmerking komen om via directe woordherkenning te worden gelezen, het proces dat als tweede werd genoemd in het voorafgaande. De rationale achter de keuze voor dit proces was dat in de normale leesontwikkeling directe woordherkenning al vlak na het prille begin een steeds prominentere plaats inneemt en de belangrijkste pijler vormt onder het automatiseringsproces. De in de inleiding genoemde stagnatie in het stadium na het prille begin zou wel eens een gevolg kunnen zijn van het niet op gang komen van het automatiseringsproces door gebreken in de directe woordherkenning. Perfetti (in press) heeft aannemelijk gemaakt dat de opbouw van het zgn. *autonome* lexicon de belangrijkste voorwaarde voor automatisering in de zin van directe woordherkenning is. Dit lexicon bevat woorden

waarvan de verschillende aspecten in sterke mate geassocieerd zijn (orthografie, klankvorm, uitspraakvorm, betekenis). Een externe stimulus in de vorm van een van deze aspecten, faciliteert direct de toegang tot de andere aspecten die, afhankelijk van de bedoeling van het verwerkingsproces, in de vorm van opslag in het werkgeheugen of in de vorm van een antwoord kunnen worden omgezet. Woorden waarbij deze associaties nog niet (of niet allemaal) direct en onherroepelijk aangeleerd zijn, bevinden zich in het zgn. *functionele* lexicon.

Een manier om te weten te komen of een bepaald woord zich in het autonome lexicon bevindt, is de stimulus op verschillende wijzen aan te bieden en het antwoord op verschillende wijzen te laten geven. Zit het in het autonome lexicon, dan hebben dergelijke variaties geen invloed op de juistheid van het antwoord. Bovendien verloopt het proces snel en vrijwel zonder dat het aandacht vraagt. Voor de ontwikkeling van het instrument heeft dit idee tot gevolg gehad dat er verschillende *aanbiedings-* en *antwoord*wijzen werden ontworpen. Wat de aanbiedingswijzen betreft, een woord kan op twee manieren aangeboden worden: visueel-orthografisch en auditief. Voor het antwoord geven zijn er meer mogelijkheden. Bij de meest gebruikelijke antwoordvorm (hardop lezen) is dat de articulatie van het woord. Ter vergelijking werd spelling opgenomen omdat Perfetti (in press) voor het meten van het autonome lexicon spelling aanbeveelt: 'Reliable, confident, and facile spelling is an index of high quality representation. By hypothesis, it is also an index of autonomous representation'(p. 29). Echter, vanwege het betrekkelijk trage (want sterk sequentiële) antwoordproces bij spelling leek het ons raadzaam om ter vergelijking nog een derde antwoordwijze te introduceren. Vergeleken met hardop benoemen gaat spellling immers veel langzamer waardoor de invloed van bewuste antwoordstrategieën (bijv. analogieregels toepassen) groter kan zijn. Bij autonome representatie wordt juist verondersteld dat dit soort strategieën niet meer toegepast behoeft te worden vanwege de automatisering van de verwerking. Besloten werd om ter vergelijking ook zogenaamde matchingstaken op te nemen die slechts ja/nee-beantwoording vergen en derhalve tijdens het antwoord geven geen aanslag hoeven te doen op bewuste antwoordstrategieën. Bij matchingstaken van twee elkaar opvolgende stimuli wordt beoordeeld of ze gelijk zijn of niet. Dit geschiedt zonder uitspraak en gaat sneller dan opschrijven. Gecombineerd met de twee aanbiedingswijzen zijn er in principe vier mogelijkheden: auditief-auditief; auditief-visueel, visueel-auditief en visueel-visueel matchen. Omdat er zich geen redenen opdrongen om a priori te selecteren, werd besloten om alle vier de mogelijkheden als subtoets in de experimentele toets op te nemen.

Een probleem als gevolg van de keuze voor onderzoek naar het autonome lexicon met verschillende aanbiedings- en antwoordwijzen was wel dat het niet mogelijk is *hetzelfde woord* op diverse manieren te laten verwerken, zonder dat er tijdens de testsessies een leereffect optreedt. Om die reden werd besloten om de analyse niet op individuele woorden uit te voeren, maar op woordklassen als geheel, waarbij de woordklasse bepaald werd door eenzelfde orthografische structuur. Kwamen leerlingen bijv. bij MKM-woorden (M = medeklinker, K = klinker) tot goede antwoorden, onafhankelijk van aanbiedings- en antwoordwijzen, dan werd die woordklasse beschouwd als opgenomen in het autonome lexicon. Daarnaast was het de vraag hoe er zeker van te zijn dat de stand van zaken in *directe woordherkenning* werd bepaald. Het is immers ook bij hoogfrequentie woorden mogelijk dat de leerling tot het goede antwoord komt door gebruik te maken van het eerste verwerkingsproces dat hij aanleert, omzetting in de klankvorm? Uit de theorie van perceptueel leren (verg. o.a. LaBerge & Samuels, 1976; Ehri & Wilce, 1983) is af te leiden dat een belangrijk aspect van het leren automatiseren de snelheid is waarmee de diverse informatieverwerkingsoperaties worden verricht. Zolang die nog traag verlopen, wordt dat als indicatie gezien van de hoeveelheid aandacht die de uitvoering kost, en daarmee van het niet-geautomatiseerd zijn. Een manier om na te gaan in hoeverre herkenningsprocessen automatisch verlopen, is de tijd te meten die het geven van een accuraat antwoord kost. Echter, antwoord geven ontwikkelt zich continu sneller tot in de volwassenheid toe (Perfetti, in press) waardoor het uitgaan van antwoordtijden op jongere leeftijd minder aantrekkelijk is. Vanwege deze ontwikkeling zou het immers moeilijk zijn om een criterium

te bepalen ter indicatie dat de leerling zich 'op de goede weg' bevindt. Een alternatief is om de aanbiedingstijden van de visueel-orthografische stimuli (woorden) te verkorten. De rationale daarachter is, dat een leerling die woorden (nog) niet op grond van woordspecifieke kennis leest, tot het goede antwoord kan komen bij een relatief lange aanbiedingstijd. Hij voert dan de omzetting in de klankvorm uit terwijl de stimulus nog zichtbaar is. Echter, wanneer de stimulus slechts korte tijd beschikbaar blijft, is het noodzakelijk om snel – en dus tot op zekere hoogte automatisch – te werken. Besloten werd om de visuele stimulus met steeds kortere aanbiedingstijd te presenteren.

Een volgende overweging betrof het gebruik van het korte termijn geheugen, hier in navolging van Baddeley (1987) *werkgeheugen* genoemd. De sequentiële aanbieding van stimuli in de matchingstaken impliceert dat het werkgeheugen een rol in de verwerking speelt. Informatie van de eerste stimulus dient vastgehouden te worden tot de tweede stimulus vergelijking mogelijk maakt. Bij spelling is er, vanwege de relatief trage aard van het antwoord geven, ook een belasting van het werkgeheugen. Hoewel er volgens Spear en Sternberg (1987) nog geen definitieve bewijsvoering is op dit punt, wordt algemeen aangenomen dat het werkgeheugen dat gebruikt maakt van een fonetische code (een klankvorm dus), bij leerlingen met ernstige leesproblemen niet goed functioneert. Dat impliceert dat, naarmate het werkgeheugen meer belast wordt, deze leerlingen meer problemen met de taak zullen hebben. Het leek ons interessant om ook in dit tekort meer inzicht te krijgen door deze factor, die toch al opgenomen was door introductie van spellings- en matchingstaken, te variëren. Dit valt te doen door een langere periode tussen stimulus en response in te lassen zodat de stimulus langer in het werkgeheugen moet worden vastgehouden. Dat gebeurde in een tweetal subtoetsen.

Ten slotte zijn er keuzen gemaakt m.b.t. de inhoud en de vorm van het instrument. Vanwege de bedoeling om het aanleren van directe woordherkenning (i.c. het opbouwen van het autonome lexicon) in kaart te kunnen brengen, is *inhoudelijk* geconcentreerd op hetgeen door de normale leerling in de loop van de tweede helft van het eerste en de eerste helft van het tweede jaar van formeel leesonderwijs wordt geleerd (resp. groep drie en vier jaar van het Basisonderwijs). Kennis van de letters is in de voorafgaande fase aangeleerd en wordt dus voorondersteld[1]. Het materiaal bestond uit een- en tweelettergrepige woorden die in de bedoelde periode verreweg de grootste proportie van de aangeboden leesinhouden vormen.

Wat de *vorm* betreft is er om drie redenen voor gekozen om het instrument computergestuurd te maken. De methodologische voordelen kunnen in een onderzoek groot geacht worden in vergelijking met traditionele toetsen door de mate van controle op input en output (Torgesen, 1986). De eerder genoemde keuzen m.b.t. aanbiedings- en antwoordwijzen maakte een grote mate van controle gewenst. Bovendien hebben we plannen om op langere termijn te komen tot een combinatie van de toets met het Computergestuurde Orthodidactische Programma voor Aanvankelijk Lezen (COPAL: Van Daal, Van der Leij, Bakker & Reitsma, 1987). Aansluiting van diagnostische informatie en instructie kan dan via één medium geregeld worden. Ten slotte, de praktische voordelen van microcomputergebruik ten behoeve van gedifferentieerde instructie van kinderen met ernstige leesproblemen zijn evident, al behoeft de software-ontwikkeling de nodige aandacht (Reitsma, 1985). Het instrument heet de Computergestuurde Toets voor Automatisering van het Lezen (COTAL).

4 *Onderzoeksopzet*

4.1 *Vraagstelling*

Nagegaan is in hoeverre de COTAL-subtoetsen verschillen aan het licht brengen tussen leerlingen met ernstige leesproblemen en leerlingen met een normale leesontwikkeling. Door proefgroepen van (kwantitatief) vergelijkbaar algemeen leesniveau te selecteren, zijn eventueel optredende verschillen, teweeggebracht door de diverse aanbiedings- en antwoordwijzen, op te vatten als indicaties voor verschillen in de kwaliteit en het ontwikkelingsverloop van het proces van directe woordherkenning.

Tabel 1 *Proefpersonen*

groep	aant.	j	m	chron.leeft.[a] gem.	s.d.	EMT[b] gem.	s.d.	dle[c] gem.	s.d.
ZZ	37	31	6	126.8	17.1	24.20	9.51	11	0.4
N	37	18	19	86.9	4.1	24.67	9.61	11	0.4
totaal	74	49	25						

a Leeftijd in maanden.
b Weergegeven is de ruwe score van de EMT
c Didactische leeftijdsequivalent (dle) is de didactische leeftijd waarop de betreffende EMT-score gemiddeld in het BaO gehaald wordt (verg. Struiksma, Van der Leij & Vieyra, 1987). Een dle van 11 wil zeggen: na een maand onderwijs in groep vier.

4.2 *Proefpersonen*

De proefpersonen met ernstige leesproblemen waren zeer zwakke lezers (ZZ) van vier LOM-scholen, de normale lezers (N) leerlingen van groep drie van vier basisscholen in de laatste maand van hun schooljaar (door de vakantiespreiding was dit eind juni/begin juli). Uit beide proefgroepen werden de leerlingen verwijderd die een EMT-score (Brus & Voeten, 1973) hadden < 10 teneinde problemen met letterkennis uit te sluiten, of > 40 om te voorkomen dat de leerlingen te goed waren. Aldus bleven 51 zeer zwakke en 66 normale lezers over. Proefpersonen in het hier gerapporteerde onderzoek zijn 37 paren van een normale en een zeer zwakke lezer, geselecteerd uit de laatstgenoemde groepen. Deze paarsgewijze matching op algemeen leesniveau was gebaseerd op vergelijkbaarheid op drie criteria: 1. EMT-ruwe score met een tolerantie voor verschil van plus of min twee punten, 2. algemene niveauscore volgens de AVI-toetsen van Van den Berg en Te Lintelo (1977) en 3. beheersing van het hardop lezen van MKM-woorden (een COTAL-subtoets die nog beschreven wordt; dit laatste criterium diende als extra controle op de onderdrempel: wanneer MKM-woorden beheerst worden, beheerst de leerling alle letters). Van matching op sekse is afgezien omdat er geen reden was om aan te nemen dat er tussen jongens en meisjes voor dit onderzoek relevante verschillen bestaan (verg. Van der Leij, 1983)[2]. Het gevolg was dat de scheve verhouding tussen jongens en meisjes die gebruikelijk is voor LOM-scholen, weerspiegeld wordt door de samenstelling van de ZZ-proefgroep (Tabel 1).

De gemiddelde leeftijd van de ZZ-groep was die van leerlingen aan het einde van groep zes van de basisschool, terwijl het leesniveau lag op begin groep vier: een gemiddelde achterstand van bijna drie jaar.

4.3 *Materiaal*

De COTAL bestaat uit subtoetsen die zijn ontwikkeld door variatie van drie dimensies: a) de aanbiedings- en antwoordwijze, b) de moeilijkheidsgraad van de aanbiedingswijze binnen de subtoets en c) de woordklasse.

a) Aanbiedings- en antwoordwijze. Er is een negental taken ontworpen waarvan er twee zijn afgevallen na de eerste toetsing[3]. In deze taken werd de stimulus (een woord) visueel of auditief aangeboden. De leerlingen antwoordden op een van de volgende drie manieren: door beoordeling van de gelijkenis met een daarop volgende stimulus ('matchen'), door natypen, of door hardop benoemen. Omdat voor een computergestuurde toets is gekozen, kreeg natypen via het toetsenbord de voorkeur boven opschrijven. Eventuele invloed van problemen met de schrijfmotoriek werd daarmee tevens uitgesloten. Tabel 2 geeft een overzicht van de subtoetsen.

b) Moeilijkheidsniveau. Het verkorten van de aanbiedingstijd van de *visuele* stimulus werd als bemoeilijkende factor opgenomen ter indicatie van de mate waarin woorden direct herkend worden. In vier subtoetsen (V1-V2, V-T, V-A, Fl.) liep het niveau van moeilijkheid op door een aanbiedingstijd van de visuele stimulus (V) van resp. 200 msec., 160 msec. en 100 msec..

Een andere bemoeilijkende conditie was het verlengen van de tijd tussen de *auditieve* stimulus en wat erop volgde (een visuele sti-

Tabel 2 *Subtoetsen en moeilijkheidsniveaus*

subtoetsen[a]		moeil. niveaus 1	2	3	betekenis niveaus
V1-V2	visueel matchen[b]	200	160	100[c]	aanb. tijd V1 en V2 neemt af
A-V	aud.-vis. match.	1''	3''	5''	tijd tussen A en V neemt toe
V-A	vis.-aud. match.[b]	200	160	100	aanb. tijd V neemt af
V-T	visueel, typen	200	160	100	aanb. tijd V neemt af
A-T	auditief, typen	1''	3''	5''	tijd tussen A en T neemt toe
Fl.	hardop lezen flitswoord	200	160	100	aanb. tijd V (Fl.) neemt af
L	hardop lezen woord	-	-	-	geen niveaus

a A = auditieve stimulus; V = visuele stimulus; T = natypen; Fl. = hardop lezen flitswoord; L = hardop lezen
b Bij V1-V2 en V-A is de tijd tussen eerste en tweede stimulus: 1.5''
c Bij de visuele stimulus is de aanbiedingstijd in msec. gegeven.

mulus om mee te matchen of de mogelijkheid tot typen: resp. A-V en A-T), ter toetsing van de invloed van belasting van het werkgeheugen. Gekozen werd voor drie niveaus: resp. met een tussentijd van 1, 3, en 5 seconden[4].

De zevende subtoets, L (woorden hardop lezen), was een gewone leestaak waarbij de stimulus op het scherm bleef staan totdat het antwoord was gegeven. Er waren bij deze subtoets dus geen moeilijkheidsniveaus.

c) Woordklassen. Alle stimuli waren hele, betekenisvolle woorden, zoveel mogelijk hoogfrequent. (Vanwege het grote aantal woorden dat voor de items nodig was, lukte het niet om uitsluitend met hoogfrequente woorden te werken; wat er aan laagfrequente woorden opgenomen werd, werd at random over de subtoetsen en moeilijkheidsniveaus verdeeld). Er werden drie soorten eenlettergrepige en drie soorten tweelettergrepige woorden opgenomen: a) MKM ('kop'), b) MKMM ('dorp'), c) MMKM ('stad'), d) twee lettergrepen: zonder regel ('keuken', 'deurmat'), e) gesloten ('lippen'), f) open ('bomen'). Vooronderstelde werd dat de moeilijkheidsgraad van deze woordklassen van a) naar f) oploopt. Deze vooronderstelling was o.a. gebaseerd op het feit dat in leesmethoden deze volgorde gehanteerd wordt en op resultaten van eerder onderzoek (verg. Van Daal et al., 1987).

Items. De woorden werden als items per aanbiedings- en antwoordwijze, moeilijkheidsniveau en woordklasse per leerling at random geselecteerd uit een bestand dat opgebouwd was bij de ontwikkeling van COPAL (Van Daal et al., 1988). Voor de matchingstaken zijn bij elk woord drie sterk gelijkende – zowel auditief als visueel – woorden gezocht. De COTAL bestond aldus uit 6 maal 6 maal 3 is 108 onderdelen (aantal aanbiedingsen antwoordwijzen X aantal woordklassen X aantal moeilijkheidsniveaus). Daarbij kwamen nog de onderdelen van de taak zonder niveaus (L): 6 woordklassen. Tesamen 114 onderdelen met 10 items elk. (Voor een vollediger beschrijving wordt verwezen naar Van Daal, Bakker, Van der Leij, Reitsma & Smeets, 1988).

4.4 *Werkwijze*

Aanbieding. De subtoetsen werden per microcomputer (en hoofdtelefoon) aangeboden (Macintosh Plus met harddisk). De visuele stimuli werden in het midden van het 17.5 X 11.5 cm grote scherm geprojecteerd met een grote, eenvoudige leesletter. Na aanbieding (200, 160 of 100 msec.) werden de visuele stimuli gemaskeerd m.b.v. een abstracte figuur die onmiddelijk, maar zeer kort, op het scherm verscheen. Dit had tot doel om de werking van het zogenaamde nabeeld op het netvlies te neutraliseren en zodoende de proefpersonen geen kans te geven om visuele kenmerken iets langer te analyseren.

De auditieve stimuli werden per hoofdtelefoon aangeboden. Het bestand aan klankvormen was verworven door de woorden op een audioband te laten inspreken door een logopediste. Door middel van een 'analog-to-digital converter' werd de natuurlijke spraak vervolgens gedigitaliseerd in de harde schijf opgeslagen.

Procedure. De subtoetsen werden stuk voor stuk afgewerkt onder leiding van een proefleider. De score was de accuratesse van de antwoorden (de instructie benadrukte dat element en niet de snelheid van antwoorden). Antwoord gaven de leerlingen door:

- het pijltje met behulp van de 'muis' naar de juiste plaats op het scherm te brengen en de toets van de 'muis' in te drukken (matchen: ja = gelijk; nee = ongelijk),
- het toetsenbord te gebruiken waarop de relevante letters waren gemarkeerd met witte plakletters en de andere dysfunctioneel gemaakt (typen),
- hardop te antwoorden.

De proefleider toetste de juistheid van dit laatste soort antwoord in en regelde het goede verloop van de toetsafname, inclusief oefentrials.

Om het beheersingsniveau gerelateerd aan het oplopen van de moeilijkheidsgraad (woordklasse; moeilijkheidsniveau) te kunnen vaststellen, werd besloten om binnen elke subtoets de onderdelen aan te bieden in een volgorde van relatief gemakkelijk naar relatief moeilijk ('bloksgewijs'). Er zijn regels geconstrueerd voor instappen, afbreken en teruggaan naar een lager niveau teneinde het afnemen van te gemakkelijke of te moeilijke onderdelen zoveel mogelijk te voorkomen. Om daarvan een voorbeeld te geven, de V1-V2-subtoets werd als eerste afgenomen met MKM-woorden en een aanbiedingstijd van resp. 200 msec., 160 en 100 msec. Daarna volgden de MKMM-woorden, etc. Beheerste de proefpersoon deze subtoets op MKMM-niveau, dan werd bij de volgende subtoets ook begonnen met MKMM-woorden. Faalde hij toch, dan werd teruggegaan naar MKM-woorden.

Proefpersonen werden niet langer dan vijf tot tien minuten per dag getoetst. Het aantal sessies per proefpersoon beliep vier à vijf dagen.

Beheersingscriterium. Een van de gevolgen van 'bloksgewijze' aanbieding was dat er per onderdeel een criterium voor beheersing moest worden gehanteerd. Besloten werd dit op 90% te stellen om een toevallige fout niet te zwaar te laten wegen. Wanneer een leerling negen van de tien items goed had werd dat onderdeel als beheerst beschouwd en kreeg de score 1. Werd dit criterium niet gehaald dan was de score 0.

De scores zijn per subtoets en woordklasse getransformeerd naar scores van 0 tot en met 5: wanneer een leerling een subtoets bij een bepaalde woordklasse beheerste op alle niveaus van moeilijkheid, kreeg hij de score 5. Over zes woordklassen heen leverde dit per subtoets een maximale score-mogelijkheid van 30 op. Ook de scores van de subtoets 'hardop lezen' die geen niveaus kent (L), zijn op deze wijze getransformeerd.

5 *Resultaten*

5.1 *Gemiddelde accuratesse per subtoets*

In Tabel 3 is de gemiddelde accuratesse voor alle subtoetsen weergegeven op twee wijzen: gesommeerd over alle woordklassen en voor MKM-woorden apart. In beide gevallen is er gesommeerd over de drie moeilijkheidsniveaus (behalve 'hardop lezen' (L) waarvoor maar één niveau bestaat).

Alle woordklassen. Geconcludeerd kan worden dat de subtoetsen de nodige moeite opleverden. De maximale score van 30 werd bij geen van de subtoetsen benaderd: er zijn geen plafond-effecten. Dat valt ook af te leiden uit het feit dat geen enkele proefpersoon alle subtoetsen beheerste (bij alle woordklassen en op alle moeilijkheidsniveaus)[5]. Wel is er een tendens naar een bodemeffect, t.w. in het geval van V-T en A-T bij N-lezers.

Over de subtoetsen heen getoetst met een multivariate variantie-analyse is het verschil tussen de twee groepen significant ($F = 7.00$, $p < .001$). De ZZ-groep was op alle matchings- en type-taken beter dan N, de N-groep was beter in hardop lezen (L), de groepen verschilden niet in flitslezen (Fl.). De verschillen bij auditief-visueel matchen (A-V), en auditief en visueel typen (A-T en V-T) waren univariaat getoetst significant. Voor visueel matchen en hardop lezen (V1-V2 en L) was er een sterke tendens tot significantie op 5% niveau.

MKM-woorden. Interessant is het om de gegevens van de MKM-woorden, de eenvoudigste woordklasse, apart te analyseren. Gegeven dat een der drie criteria die als onderdrempel fungeerden in de selectie, de beheersing betrof van het hardop lezen (L) van MKM-woorden, zou het immers kunnen zijn dat ook andere subtoetsen bij deze woordklasse door de leerlingen beheerst werden.

Tabel 3 *Gemiddelden en standaarddeviaties van ZZ en N per subtoets, gesommeerd over drie moeilijkheidsniveaus, resp. voor alle woordklassen en MKM-woorden*

subtoets	alle woordklassen[a]				MKM-woorden			
	ZZ	N	F	p	ZZ	N	F	p
V1-V2	7.2	3.4	3.45	n.s.[b]	2.6	0.8	20.59	<.001
	(8.1)	(7.8)			(1.8)	(1.5)		
A-V	17.2	10.8	7.23	<.01	3.7	0.8	11.82	<.001
	(8.6)	(12.2)			(1.6)	(1.5)		
V-A	8.5	5.5	1.98	n.s.	3.3	1.7	13.00	<.001
	(7.8)	(8.2)			(1.7)	(2.0)		
V-T	4.9	1.5	7.16	<.01	2.2	0.9	9.31	<.01
	(6.4)	(3.3)			(1.9)	(1.7)		
A-T	10.8	2.1	22.76	<.001	3.7	0.8	66.36	<.001
	(9.6)	(5.0)			(1.6)	(1.5)		
Fl.	6.0	5.9	0.00	n.s.	2.8	2.1	1.78	n.s.
	(7.1)	(7.9)			(1.8)	(2.2)		
L[c]	13.0	16.8	3.78	n.s.[d]	–[e]	–	–	–
	(8.9)	(7.8)						
Max. sc.	30	30			5	5		

a Standaarddeviaties tussen haakjes
b p = .064
c Deze subtoets kent geen moeilijkheidsniveaus.
d p = .056
e Het hardop lezen (L) van MKM-woorden werd als onderdrempel in de selectie gebruikt. Alle 74 leerlingen beheersten dit (score 5.0).

Tabel 3 toont echter dat dit bij geen van de andere subtoetsen voor de groepen als geheel het geval was. Nadere analyse leert, dat slechts vier van de 37 ZZ-lezers en niet één N-lezer alle subtoetsen bij MKM-woorden op alle niveaus van moeilijkheid beheersten.

De verschillen tussen ZZ en N waren bij MKM-woorden nog opvallender dan bij alle woordklassen tesamen (multivariaat getoetst: $F = 15.06$, $p < .001$). ZZ was significant beter in de matchings- en type-taken (voor de univariate toetsing, zie Tabel 3); het verschil in flitslezen was echter gering en niet significant.

5.2 *Analyse van kwalitatieve verschillen*

Teneinde na te gaan in hoeverre de verschillende aanbiedings- en antwoordwijzen, moeilijkheidsniveaus en woordklassen significante verschillen te zien geven tussen de groepen is een aantal multivariate interactie-analyses uitgevoerd (MANOVA's) met deze factoren als binnenproefpersoonvariabelen en ZZ/N als tussenproefpersoonvariabele. Voor de interpretatie is voorts gebruik gemaakt van univariate F's en de gemiddelde scores (zie o.a. Tabel 3). Omdat uit de gegevens van Tabel 3 is op te maken dat de beide typetaken voor de N-groep te moeilijk waren (op dit punt wordt teruggekomen), zijn ze uit de analyse naar de invloed van antwoordwijze weggelaten.

Aanbiedingswijze. De interactie tussen ZZ/N en aanbiedingswijze (eerst auditief of visueel, twee niveaus) is onderzocht door A-V te contrasteren met V-A en bleek significant (F (df 2, 70) = 3.57; $p = .03$). Uit Tabel 3 valt op te maken dat dit komt omdat de ZZ-lezers relatief meer profiteren van de (eerst) auditieve aanbieding dan de N-lezers (het verschil was bij A-V significant, bij V-A niet).

Antwoordwijze. Over alle woordklassen gerekend, is de interactie tussen ZZ/N en matchen (V1-V2 en V-A gesommeerd) versus hardop benoemen (L en Fl. gesommeerd), significant (F (df 2, 70) = 8.87, $p < .0005$). Uit Tabel 3 is af te leiden dat dit enerzijds komt door een bijna significant verschil ten gunste van ZZ bij de V1-V2 subtoets, anderzijds door een bijna significant verschil ten gunste van N bij de L-subtoets.

Moeilijkheidsniveaus. Verkorting van visuele stimulus interacteert met ZZ/N wanneer de onverkorte aanbieding (L) en Flitslezen bij 200, 160 en 100 msec. (vier niveaus) in de analyse worden betrokken (F (df 4, 69) = 4.32;

$p < .005$). Wanneer de berekening wordt gereduceerd tot onverkort aanbieden/aanbieding van 200 msec. (twee niveaus), blijft de interactie significant (F (df 2, 71) = 3.97; $p = .02$). Op het niveau van 160 en 100 msec. is dat niet het geval. Gegeven het op 5%niveau bijna significante verschil ten gunste van N op de L-subtoets (Tabel 3) indiceert dit dat normale lezers relatief meer moeite hebben met verkorting van de stimulus tot 200 msec. dan de zeer zwakke lezers.

Verlenging van de tijd tussen de auditieve stimulus en de visuele stimulus (A-V, drie niveaus: 1'', 3'', 5'') interacteert ook met ZZ/N (F (df (3,70) = 11.76; $p = .0001$), vooral ten gevolge van een relatief betere score van ZZ bij op het 1'' niveau ($F = 15.04$; $p = .0002$) en 3'' niveau ($F = 5.92$; $p < .02$). Bij 5'' zijn er geen significante verschillen tussen de groepen. De hogere score van ZZ bij A-V is dus niet toe te schrijven aan een surplus in de werking van het werkgeheugen – dan zou het verschil er ook zijn bij 5'' – maar aan het profijt dat ze trekken van auditieve aanbieding op het niveau dat het werkgeheugen minder belast wordt.

Woordklasse. Woordklasse (zes niveaus) interacteert met ZZ/N (F (df 6, 67) = 9.65; $p = .0001$), geheel toe te schrijven aan de relatief betere score van ZZ bij de eenvoudigste woorden (MKM) ten opzichte van gelijkwaardiger scores bij de moeilijker woordklassen. Deze uitspraak geldt echter wanneer alle subtoetsscores worden ingecalculeerd. Wanneer de twee subtoetsen (A-V, L) geselecteerd worden die, blijkens het voorafgaande, de duidelijkste verschillen te zien geven, is er een significante driewegsinteractie tussen ZZ/N, woordklasse en A-V/L. Daarbij zijn de woordklassen teruggebracht tot drie niveaus (MKM, MKMM + MMKM, tweelettergrepige woorden): F (df 5,67) = 3.87; $p = .004$. Uit de univariate F's en de gemiddelde scores valt op te maken dat zeer zwakke lezers met de A-V-subtoets significant beter zijn dan N wanneer het om eenlettergrepige woorden gaat (MKM en MKMM + MMKM; resp. $F = 11.82$; $p = .001$; $F = 4.12$; $p < .05$) maar niet bij tweelettergrepige woorden. Normale lezers zijn daarentegen relatief beter in het hardop lezen van MKMM + MMKM-woorden ($F = 5.95$; $p = .02$), maar niet van tweelettergrepige woorden. Bij het hardop lezen van MKM-woorden waren de groepen, ten gevolge van de matchingsprocedure, gelijk. Waarschijnlijk zijn tweelettergrepige woorden voor beide groepen te moeilijk. Belangrijker is echter de constatering dat wanneer er eenlettergrepige woorden van MKMM- of MMKM-structuur worden aangeboden, zeer zwakke lezers significant beter zijn dan normale wanneer eerst een auditieve stimulus aangeboden wordt, te matchen met een visuele. Voor hardop lezen geldt juist dat normale lezers bij MKMM- en MMKM-woorden significant beter zijn. In de discussie wordt hierop teruggekomen.

5.3 *Mate van beheersing*

Tot nu toe betreffen de analyses de gemiddelde scores van de groepen. Hoewel de data interessante verschillen tonen, wordt daarmee geen inzicht verschaft in de mate waarin de leerlingen de subtoetsen op de diverse moeilijkheidsniveaus beheersen. Aangezien er criteriumtoetsen met afbreekregels gebruikt zijn, is ook dat te demonstreren, niet door de gemiddelde scores maar door de aantallen leerlingen die tot beheersing komen. Het is niet doenlijk om in het kader van een artikel een compleet overzicht hiervan te geven. Ter illustratie staan in Tabel 4 de aantallen leerlingen weergegeven die de subtoetsen beheersten waarbij hardop moest worden gelezen, apart per woordklasse (voor Fl. en L) en moeilijkheidsniveaus (voor Fl.).

Behalve bij het hardop lezen van MKM-woorden (L) werden de L- en Fl.-subtoetsen niet beheerst door alle leerlingen. Uit Tabel 4 blijkt dat er een groot verval was in aantal leerlingen dat de subtoetsen beheerste naarmate 1. de orthografische structuur van de woordklasse gecompliceerder is, en 2. het moeilijkheidsniveau van de visuele stimulusaanbieding oploopt.

De eerste tendens, het *woordklasse-effect*, is geleidelijker bij de normale leerlingen, dan bij de zeer zwakke lezers die vooral met het hardop lezen van MKMM- en MMKM-woorden relatief meer moeite hadden. Dit grotere verval bij ZZ is in zekere mate ook terug te vinden in de scores op de flitsleestaak (Tabel 4) en de andere taken (niet weergegeven) en verklaart de bij woordklasse genoemde interacties. Dit is ook de reden dat het verschil tussen ZZ en N in Tabel 3 groter is bij MKM-woorden alleen dan bij alle woordklassen. Omgekeerd kan men ook zeggen dat ZZ-

Tabel 4 *Aantallen leerlingen die de L- en Fl.subtoetsen beheersen, per woordklasse en moeilijkheidsniveau (max. = 37)*

taak:	L[a]		Fl.[200]		Fl.[160]		Fl.[100]	
woordklasse	ZZ	N	ZZ	N	ZZ	N	ZZ	N
MKM[b]	-	-	32	21	20	16	10	10
MKMM	22	31	17	15	9	12	5	9
MMKM	15	27	9	12	7	9	2	5
2 l.gr.	10	14	6	5	3	4	0	3
gesl. l.gr.	9	12	5	4	4	4	2	2
open l.gr.	3	3	2	1	1	1	1	0

a Deze subtoets kent geen moeilijkheidsniveaus.

b Het hardop lezen (L) van MKM-woorden werd als onderdrempel in de selectie gebruikt. Alle 74 leerlingen beheersten dit.

lezers beter waren dan N-lezers in het verwerken van MKM-woorden mits de response geen articulatie vereiste.

De tweede tendens, het *flitsaanbiedingseffect*, is juist geleidelijker bij ZZ wanneer L (moeilijkheidsgraad nul vanwege onbeperkte aanbiedingstijd) wordt vergeleken met de reeks Fl.[200] tot Fl.[100]. Voor alle woordklassen geldt dat de verkorting van de aanbiedingstijd tot 200 msec., vergeleken met een onbepaalde aanbiedingstijd, meer N-leerlingen moeite gaf dan de ZZ-leerlingen, conform de bij moeilijkheidsniveaus eerder genoemde interactie. Omgekeerd kan men ook zeggen dat de N-lezers beter waren dan ZZ-lezers in het hardop lezen van woorden, die zichtbaar bleven.

Tabel 4 demonstreert behalve een ondersteuning van de interacties, nog twee punten. Ten eerste, de groepen van leerlingen verdelen zich keurig over de beheersingsniveaus van deze twee subtoetsen: het eenvoudigste niveau werd door allen beheerst en het moeilijkste door slechts 1 (uit de ZZ-groep). Het oplopen van het moeilijkheidsniveau door verkorting van de visuele aanbieding en aanbieding van moeilijker woorden doet de een na de ander afvallen. Daarmee is het diagnostisch differentiërende aspect van deze subtoetsen in kwantitatieve zin (van meer naar minder beheersing) evident. Ten tweede, inspectie van Tabel 4 leert ook dat de beheersingspatronen van ZZ en N over het geheel genomen, verschillen. Alle leerlingen beheersten het hardop lezen van MKM-woorden. Echter, van de 37 ZZ-leerlingen deden 32 dat ook bij Fl.[200], maar slechts 22 bij het hardop lezen van MKMM-woorden. Van de 37 N-leerlingen beheersten er slechts 21 MKM-woorden bij Fl.[200], maar 31 het hardop lezen van MKMM. Deze verschillen zijn er resp. ten gevolge van het woordklasse- en het flitsaanbiedingseffect. Combinatie met andere subtoetsen (A-V, V1-V2) ondersteunt het idee dat het instrument kwalitatieve verschillen in beheersingspatroon tussen ZZ en N indiceert, niet alleen als groep maar ook individueel.

6 *Discussie en conclusies*

6.1 *Kwantitatieve analyse*

Met betrekking tot de eerste bedoeling van het instrument – instructiegevoeligheid indiceren, uitgedrukt in beheersingsniveaus van criteriumtoetsen (zie de inleiding) – kan als voorlopige conclusie getrokken worden dat het instrument in *kwantitatief* opzicht bevredigende resultaten te zien geeft. De subtoetsen waren gemiddeld bepaald niet te gemakkelijk en er was een grote spreiding van scores van proefpersonen op een gemiddeld algemeen leesniveau van begin groep vier. Dat er geen plafondeffect is blijkt bijvoorbeeld uit de spreiding die Tabel 4 te zien geeft. Wordt de eenvoudigste taak (het hardop lezen van MKM-woorden) nog door 100% van de leerlingen goed gedaan – een gevolg van de selectieprocedure – het aanbieden van moeilijker woordklassen en kortere stimulustijden reduceert deze maximale proportie vrijwel tot nul. Sommige subtoetsen zijn echter aan de moeilijke kant: de beide type-taken leverden bij de N-groep lage scores op, ook bij MKM-woorden. Vanwege dit bodemeffect zijn deze subtoetsen niet in de interactie-analyses betrokken. Verondersteld kan worden dat de

ontwikkeling van de processen die betrokken zijn in deze taken in het normale geval plaatsvindt in de periode na de toetsperiode (eind groep drie) en dat de invloed van instructie daar groot op is. Dat betekent dat op den duur het geconstateerde verschil aan betekenis verliest, omdat het niet te herleiden is tot specifieke tekorten (bij N) of compensaties (bij ZZ) in onderliggende verwerkingsprocessen, maar tot een gebrek aan ervaring van de normale lezers. Ook op grond van de literatuur valt niet te verwachten dat zeer zwakke lezers superieur zouden zijn in spelling door toedoen van een surplus in capaciteiten (zie bijv. Seymour, 1986). Ze hebben in het algemeen met spelling (minstens) evenveel moeite als met hardop lezen.

Duidelijk is voorts dat de proefpersonen, op een paar uitzonderingen na, zelfs de eenvoudigste woorden (MKM) niet bij alle aanbiedings- en antwoordwijzen beheersten. Comform de ideeën van Perfetti (in press) indiceert dit dat ook deze woordklasse zich nog in het stadium van het functionele lexicon bevindt. Wanneer het ontwikkelingsproces normaal verloopt, valt te verwachten dat in de loop van de leerjaren die volgen op de toetsperiode (eind groep drie), de woordklassen een voor een in het autonome lexicon terecht komen. Dit is af te leiden uit het feit dat de woordklassen beheerst worden, onafhankelijk van aanbiedingsantwoordwijze en moeilijkheidsniveau. Het zal pas mogelijk zijn om een uitspraak te doen over de validiteit van de operationalisering van het begrip autonoom lexicon in de vorm van het instrument, wanneer deze verwachting uit blijkt te komen. Dit vereist vervolgonderzoek bij dezelfde leerlingen (zie de afsluitende paragraaf).

6.2 *Kwalitatieve analyse*

De tweede bedoeling van het instrument, indicatie van *kwalitatieve* verschillen tussen zeer zwakke en normale lezers van gelijk leesniveau, vormde de hoofdvraag van het gerapporteerde onderzoek. Uit de eerste analyse kan worden afgeleid dat 1. de ZZ-groep beter was in auditief-visueel matchen (A-V), 2. hardop lezen (L) door de normale lezers beter werd gedaan, 3. een grotere proportie van de ZZ-groep moeite had met het oplopen van de orthografische moeilijkheid van de (eenlettergrepige) woordklassen, en 4. een groter deel van de N-groep moeite had met het oplopen van de moeilijkheid door verkorting van de aanbiedingstijd (de overgang van onbeperkte aanbieding naar msec.). Uit deze opsomming kan geconcludeerd worden dat het variëren van aanbiedings- en antwoordcondities, moeilijkheidsniveaus en woordklassen verschillen in beheersingspatronen tussen zeer zwakke en normale lezers van hetzelfde leesniveau aan het licht brengt. In het navolgende zal geen poging gedaan worden deze gegevens te integreren tot een differentiële leestheorie voor normale en zeer zwakke lezers. Ontwikkeling van de beheersing van de leesprocessen, geïndiceerd door het vervolgonderzoek, zal daarover meer uitsluitsel kunnen geven. We beperken ons tot een specificatie en een voorlopige interpretatie van de gevonden verschillen.

Ad 1. Wat de A-V-taak betreft is het niet waarschijnlijk dat een verschil in ervaring het verschil bepaalt, zoals bij de type-taken het geval is: een dergelijke taak komt in het onderwijs niet of slechts sporadisch voor. Het lijkt er meer op dat de auditieve aanbieding van de articulatorische vorm woordherkenning meer *faciliteert* bij zeer zwakke lezers dan bij normale lezers van gelijk leesniveau. De moeite die een behoorlijk groot deel van de ZZ-leerlingen blijkt te hebben met de overgang van MKM- naar MKMM- en MMKM-woorden (zie ad 3.) is in deze taak ook veel minder evident. Ter illustratie: op de A-Vsubtoets (niveau 1) beheersten 37 van de ZZ-leerlingen de MKM-woorden, 33 de MKMM-woorden en 32 de MMKM-woorden. Ter vergelijking: de aantallen die genoemd werden in Tabel 4 m.b.t. de hardop leestaak (L) zijn resp. 37, 22 en 15.

Het verschil bij A-V ten gunste van ZZ kan niet herleid worden uit een surplus in de werking van het *werkgeheugen* bij de zeer zwakke lezers, zoals de interactie-analyse liet zien. Wanneer er meer aanslag op dat geheugen wordt gedaan omdat de informatie 5'' moet worden vastgehouden, verdwijnen de verschillen immers. Het lijkt erop dat, zodra het werkgeheugen langduriger belast wordt er tussen de 3'' en 5'' aanbieding een kritische grens wordt gepasseerd. De informatie in het werkgeheugen is niet meer zo duidelijk aanwezig en het verwerken van de visuele stimulus (dat veel aandacht kost) interfereert daarmee. In het vervolgonderzoek zal nagegaan worden in hoeverre dit een stabiel kenmerk

van zeer zwakke lezers is dat afwijkt van het normale ontwikkelingspatroon. Dan zal het mogelijk zijn om enig licht te werpen op de eerder genoemde meningen van Baddely (1987) en Spear & Sternberg (1987), die op dit punt contrasteren[6].

Ad 2. Aangezien de leerlingen paarsgewijs gematched waren op het hardop lezen (op drie criteria: EMT, AVI, beheersing hardop lezen van MKM-woorden), is het resultaat ten gunste van de normale lezers bij het *hardop lezen* (L) opmerkelijk te noemen. Het moge echter duidelijk zijn dat de EMT en AVI niet identiek zijn aan L. Bij de EMT gaat het om het lezen van woorden in rijtjes, bij AVI om woordjes in zinnen. Snelheid speelt in beide toetsen een rol: bij de EMT is de score het aantal goed gelezen woorden in één minuut en bij AVI is er, per kaart, een beheersingscriterium voor de snelheid (en voor de accuratesse). De hardop lees-toets L betreft daarentegen het lezen van stuk voor stuk aangeboden woorden. Het snelheidsaspect speelt in de score geen rol, het gaat alleen om het behalen van het 90% goed criterium per woordklasse. In de EMT en AVI zijn bovendien allerlei processen van invloed die niet uitsluitend met het lezen van woorden te maken hebben: o.a. het verspringen van regel (vooral bij de EMT) en het gebruik maken van de context (AVI). Men kan L beschouwen als een zuiverder toets voor het lezen van losse woorden dan EMT en AVI.

Univariate toetsing over de woordklassen heen gaf bij L een marginaal significant verschil te zien (Tabel 3; $p = .056$). Uit nadere analyse per woordklasse bleek echter dat de verschillen bij MKMM- en MMKM-woorden wel significant waren (verg. ook de aantallen leerlingen die tot beheersing kwamen in Tabel 4.) Voor de interpretatie van het resultaat is het van belang te benadrukken dat, in tegenstelling tot de matchingssubtoetsen, in de hardop leestoets articulatie betrokken is. Gewezen kan worden op de parallel met de problemen die zeer zwakke lezers hebben met taken die het benoemen van overbekende stimuli vereisen (bijv. van cijfers, letters of plaatjes). Baddeley's (1987) conclusie dat het omzetten in de articulatorische code als onderdeel van het produktieproces bij slechte lezers een zwak punt is, krijgt door ons resultaat ondersteuning (verg. ook Spring & Capps, 1974; Spring & Perry, 1983). Echter, Baddeley beschrijft de problemen als snelheidsproblemen. Aangezien snelheid van articulatie niet de score was waarmee we werkten, indiceert het resultaat met de L-subtoets o.i. dat zeer zwakke lezers ook met de *accurate produktie* van de articulatorische code meer moeite hebben dan normale lezers van overeenkomstig leesniveau. Dat het een produktie- en geen herkenningsprobleem is ontlenen we tevens aan vergelijking met de A-V-taak (zie ad 1). Aangezien in beide subtoetsen de visuele stimulus onverkort aangeboden werd, spitst het verschil in taakeisen zich toe op herkenning en opslag van de auditief aangeboden articulatievorm (A-V) versus produktie van die vorm (L). Zeer zwakke lezers waren met auditief-visueel matchen bij alle woordklassen beter dan met hardop lezen (resp. gem. 17.2 en 13.0), in contrast met N die voor A-V en L resp. gem. 10.8 en 16.8 behaalden (zie Tabel 3). Dit articulatorische produktieprobleem is een belangrijk kenmerk van het leesproces van zeer zwakke lezers.

Ad 3. De sprong van MKM- naar MKMM- en MMKM-woorden vormde voor een groter deel van de ZZ-lezers een probleem dan voor N-lezers (het *woordklasse-effect*). Voor de hand ligt om dit toe te schrijven aan het feit dat ZZ-lezers de MKM-woorden significant accurater verwerkten in de matchingstaken, onafhankelijk van de aanbiedingswijze (V1-V2, A-V, V-A). Dit surplus valt – op A-V na – weg wanneer alle woordklassen ingecalculeerd worden (zie Tabel 3). (Bij L kan er geen verschil optreden bij MKM-woorden vanwege de selectieprocedure; het niet-significante verschil bij Fl. – ook een hardop leestaak – is daar waarschijnlijk een gevolg van). Echter, voor de A-V-subtoets geldt dat de zeer zwakke lezers ook MKMM- en MMKM-woorden significant beter verwerkten dan normale lezers, maar voor hardop lezen is het omgekeerde het geval. Dit verschil kan te wijten zijn aan het tekort in het produceren van de articulatorische code dat ad 2. als kenmerk van zeer zwakke lezers genoemd werd. Het produktieprobleem is in de A-V-taak sterk gereduceerd: er wordt een uitspraakvorm gegeven (A) die moet worden onthouden en vergeleken met een orthografisch patroon (V). Anders gezegd: de auditief aangeboden uitspraakvorm compenseert het tekort door reductie van het aantal alternatieven in de 'inner speech store' waaruit gekozen moet worden tot één (een 'priming' effect).

Bij L is een dergelijke compensatie er niet: de leerling is volkomen afhankelijk van zijn vermogen om de juiste uitspraakvorm te genereren. Deze interpretatie impliceert echter dat ook in een matchingstaak als A-V de articulatorische code een rol speelt. Als theoretisch argument dat deze interpretatie ondersteunt kan gesteld worden dat het belang van de articulatorische code bij het leren lezen door vele auteurs groot wordt geacht (verg. Span, 1982, voor een overzicht). Het is echter op grond van onze gegevens niet te concluderen dat het verschil tussen zeer zwakke en normale lezers uitsluitend het gevolg is van het voornoemde tekort bij de zwakke lezers. Dat wordt geïllustreerd door het navolgende.

Ad 4. Verkorting van de visuele stimulus van onbeperkte aanbieding tot 200 msec. leidt er bij normale lezers toe dat ze hun relatieve voorsprong op zeer zwakke lezers bij hardop lezen verliezen (het *flitsaanbiedingseffect*). Dit geldt overigens alleen bij eenlettergrepige woorden – bij tweelettergrepige woorden zijn de verschillen bij onverkorte aanbieding gering omdat er in beide groepen veel leerlingen zijn die het hardop lezen niet op 90% niveau beheersen. Dit indiceert dat een behoorlijk proportie van de normale lezers voor de associatie van het visueel-orthografische patroon aan de uitspraakvorm nog vrij veel tijd nodig heeft. Meer dan de helft van degenen die het hardop lezen van MKMM- en MMKM-woorden nog beheersten, deden dat niet bij verkorting tot 200 msec. (zie Tabel 4). Het lijkt er op dat hun woordspecifieke kennis (nog) sterk gebonden is aan de mogelijkheid om het herkennen van de visueel-orthografische stimulus en het produceren van de uitspraakvorm simultaan uit te voeren. Voor zeer zwakke lezers geldt dat een groot aantal dit bij het hardop lezen van MKMM- en MMKM-woorden niet beheerste (resp. 15 en 22 van de 37; Tabel 4). Omdat het voor het overgrote deel om hoogfrequente woorden ging, kan aangenomen worden dat de woordspecifieke kennis van deze leerlingen geringer is, in overeenstemming met hetgeen Reitsma (1983) concludeert. Hoogfrequente woorden komen immers door veelvuldig gebruik in spreek- en schrijftaal het eerst in aanmerking om middels woordspecifieke kennis aangeleerd te worden.

De relatie tussen de beide genoemde verschillen tussen normale en zeer zwakke lezers, in het aanleren van woordspecifieke kennis en het produceren van de articulatorische code, is een interessant theoretisch vraagstuk waar we in het kader van dit artikel niet op in kunnen gaan, op enkele suggesties na. Aangenomen dat de N-groep model staat voor een normale ontwikkeling, stellen we op grond van onze resultaten als hypothese dat de associatie van de orthografische structuur met de uitspraakvorm de kern vormt van woordspecifieke kennis en daarmee van de opbouw van het autonome lexicon. Dergelijke associaties leggen kost zeer zwakke lezers moeite, waarschijnlijk ten gevolge van een tekort in het produceren van de uitspraakvorm. Het eerst beschikken over de auditief aangeboden articulatievorm vergemakkelijkt woordherkenning bij zeer zwakke lezers: de noodzaak tot het zelf genereren van de uitspraakvorm wordt van buitenaf gereduceerd. Echter, dit is geen 'lezen' in de gebruikelijke zin van het woord: de visuele stimulus is er dan altijd eerst en er moet een antwoord geproduceerd worden. Wel kan er in therapeutisch-compenserende zin van dit kenmerk van zeer zwakke lezers gebruik gemaakt worden (zie de afsluitende paragraaf).

6.3 *Implicaties voor verder onderzoek*

Uit het voorafgaande moge duidelijk geworden zijn dat de analyse van de eerste gegevens van het onderzoeksproject een aantal interessante veronderstellingen heeft opgeleverd over de eigenaardigheden van ernstige leesproblemen die met COTAL kunnen worden geduid. Echter, het instrument (of een aangepaste versie daarvan) zal pas getoetst kunnen worden op differentiële validiteit in vervolgonderzoek. Zo is het noodzakelijk om veronderstelde verschillen in het ontwikkelingsverloop van het automatiseringsproces tussen normale en zeer zwakke lezers, aan te tonen. Inmiddels is de beschreven toetsronde gevolgd door nog twee toetsronden bij de N-groep. Aangenomen dat het automatiseringsproces in die periode in volle gang is, verwachten we dat in de loop van groep vier (om precies te zijn in november/december en in maart/april, ongeveer vijf en negen maanden na de eerste afname) de N-groep tot beheersing van steeds meer onderdelen komt en blijk geeft van een zich uitbreidend autonoom lexicon. Beheersing van een woordklasse (i.c. beheersing van de subtoetsen op alle moeilijk-

heidsniveaus binnen die klasse) zal, gegeven de opzet van het instrument, de indicatie zijn voor opname van die klasse in het autonome lexicon.

Ook de zeer zwakke lezers zijn opnieuw, zij het met een langere tussenpoos, getoetst. De vraag die o.i. van groot belang is voor onderscheid tussen normale en zeer zwakke lezers is in hoeverre zeer zwakke lezers in staat zijn om tot een autonoom lexicon te komen, dat wil zeggen welke grenzen daar in termen van woordklassen, moeilijkheidsniveaus en aanbiedings- en antwoordwijzen aan te ontdekken zijn. Naar verwachting zal de vooruitgang bij een groot aantal van de zeer zwakke lezers zo gering zijn in vergelijking met die van normale lezers, dat van een autonoom lexicon nauwelijks gesproken kan worden. Uiteraard zullen in de analyse van de longitudinale data de eerder geventileerde ideeën omtrent de verschillen in ervaring (typen), in woordspecifieke kennis en in produktie van de articulatorische code, aan een nadere toetsing onderworpen worden.

Een volgende stap is onderzoek doen naar de vraag of uit de gegevens de instructiebehoeften van zeer zwak lezende leerlingen zijn af te leiden: de condities waaronder het effect van instructie zo groot mogelijk zal zijn. Het lijkt voorlopig duidelijk dat auditieve aanbieding van de articulatievorm een essentieel compenserend instructieprincipe is. Onderzoek waarin dit op verschillende wijze is getoetst, ondersteunt deze uitspraak (verg. o.a. Van der Leij, 1983, 1986; Bouwhuis & Truin, 1986; Van der Leij & Van Daal, 1989; Van Daal & Reitsma, 1989). De belangrijkste vraag is uiteraard in hoeverre dergelijke instructie leidt tot 'autonomiseren' van het lexicon en daarmee tot automatiseren van het lezen.

Ten slotte wijzen we op de relatie tussen de toets en instructieprogramma's i.c. tussen COTAL en het Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen (COPAL) waarvan auditieve aanbieding een belangrijk onderdeel vormt (Van Daal et al., 1987). Een effectiviteitsonderzoek waarbij toets en programma gecombineerd worden staat als sluitstuk van het project op het programma.

Noten

1. Men kan zich afvragen in hoeverre deze vooronderstelling gerechtvaardigd is. Uit (nog ongepubliceerd) vervolgonderzoek blijkt dat leerlingen met een vergelijkbare leeszwakte als de proefpersonen die in de onderhavige studie zijn onderzocht, zich niet onderscheiden van normale lezers die gematched zijn op leesniveau, in accuratesse en snelheid van het benoemen van cijfers en letters. Ook herkenning of een stimulus een letter of een cijfer is gaf geen verschil te zien. Mocht de letterkennis nog niet geautomatiseerd zijn op dit leesniveau (eind groep drie/begin groep vier) dan zijn normale en zeer zwakke lezers in dit opzicht vergelijkbaar.
2. Een andere mogelijke matchingsvariabele, intelligentie, werd niet gebruikt omdat we aannamen dat er door het toelatingsbeleid van LOM-scholen – slechts normaal intelligente kinderen worden aangenomen – en door het verwijderen van BaO-leerlingen die te goed lazen (zeker zo'n 25% van de oorspronkelijke steekproef), weinig verschil in zou zijn. Het is echter mogelijk dat deze assumptie niet geheel gerechtvaardigd is. Om die reden zal in het follow-up gedeelte van dit longitudinale onderzoek alsnog een vorm van intelligentiemeting uitgevoerd worden. Mochten er significante verschillen optreden dat kan die meting hetzij tot correctie van de gematchte groepen, hetzij tot invoering van intelligentie als covariaat leiden.
3. In een eerdere versie van de toetsbatterij waren nog twee subtoetsen opgenomen: auditief matchen en nazeggen van woorden. De eerste subtoets (A-A) maakte het kwartet matchingsmogelijkheden tussen auditieve en visuele stimuli compleet. Analoog aan de A-T- en A-V-subtoets werd de moeilijkheidsgraad verhoogd door tussen de stimuli het tijdsverloop te laten oplopen van 1 naar 3 naar 5 seconden. De tweede subtoets, nazeggen, had ten doel om na te gaan in hoeverre uitspraak op zichzelf problemen gaf. Bij 51 LOM-leerlingen die deze versie kregen bleek het percentage dat deze subtoetsen accuraat aflegde over alle woordklassen en moeilijkheidsniveaus heen, 95 te bedragen. Vanwege het niet-differentiërende karakter van deze subtoetsen waar het de groep van lezers betrof waarin we het meest geïnteresseerd waren, zijn ze in het verdere onderzoek weggelaten.
4. In een eerdere versie waren nog twee tussenliggende moeilijkheidsniveaus opgenomen: 180 en 130 msec. (V1-V2, V-A, V-T, Fl.) en 2" en 4" (A-V, AT). Omdat de scores bij het oplopen van het moeilijkheidsniveau bij alle woordklassen en subtoetsen lineair verliepen, zijn deze tussenliggende waarden weggelaten in het verdere onderzoek.

5. Vier ZZ-leerlingen beheersten één subtoets bij alle woordklassen, één ZZ-leerling beheerste er twee; vijf N-leerlingen beheersten er één. Van deze in totaal elf beheerste subtoetsen ging het in zes gevallen om L, de enige taak zonder moeilijkheidsniveaus. Vier keer betrof het A-V en één keer V1-V2, subtoetsen met drie moeilijkheidsniveaus.
6. Overigens is er, op basis van de gepresenteerde gegevens, geen uitspraak mogelijk over de kwestie in hoeverre de werking van het werkgeheugen bij zeer zwakke lezers afwijkt van die van leeftijdgenoten om dat het hier gaat om de vergelijking met jongere leerlingen van gelijk leesniveau. In het follow-up onderzoek zal deze vergelijking wel worden gemaakt.

Literatuur

Aarle, E.J.M. van, *Behandeling van ernstige leesproblemen.* Een vergelijkend onderzoek naar de effectiviteit van zelfinstructie en training in decoderen en contextgebruik. Nijmegen: 1988 (diss.).

Baddeley, A., *Working memory.* Oxford: Clarendon Press, 1987.

Berg, R.M. van den & H.G. te Lintelo, *A.V.I.-pakket.* Den Bosch: K.P.C., 1977.

Bouma, H. & C.P. Legein, Foveal and parafoveal recognition of letters and words bij dyslectics and by average readers. *Neuropsychologica,* 1977, *15,* 69-80.

Bouma, H. & C.P. Legein, Dyslexia: a specific recoding deficiency. An analysis of response latencies for letters and words in dyslectics and average readers. *Neuropsychologica,* 1980, *18,* 285-298.

Bouwhuis, D.G. & P.G.M. Truin, Woorden lezen met spraak naar keuze door dyslectische kinderen. In: S. Dijkstra & P. Span (Red.), *Leerprocessen en instructie.* Lisse: Swets & Zeitlinger, 1986.

Brus, B.Th. & M.J.M. Voeten, *Een-minuuttest, vorm A en B.* Nijmegen: Berkhout, 1973.

Daal, V.H.P., A. van der Leij, N.C.M. Bakker & P. Reitsma, Een Computergestuurd Programma voor Aanvankelijk Lezen. *Pedagogische Studiën,* 1987, *64,* 364-376.

Daal, V.H.P., N.C.M. Bakker, A. van der Leij, P. Reitsma & H. Smeets, COPAL: *Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen.* Verslag over de ontwikkeling en beproeving. Amsterdam: Vrije Universiteit & Paedologisch Instituut, 1988.

Daal, V. van & P. Reitsma, Audio support in the remediation of reading problems. In: J.J. Dumont & H. Nakken (Eds.), *Learning desabilities. Volume 2. Cognitive, social and remedial aspects.* Lisse: Swets & Zeitlinger, 1989, 149-158.

Ehri, L.C. & L.S. Wilce, Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology,* 1983, *75,* 3-18.

LaBerge, D. & S.J. Samuels, Toward a theory of automatic information processing in reading. In: H. Singer & R.B. Ruddell (Eds.), *Theoretical models and processes of reading* . Newark: IRA, 1976.

Leij, A. van der, *Ernstige leesproblemen.* Een onderzoek naar mogelijkheden tot differentiatie en behandeling. Lisse: Swets en Zeitlinger, 1983 (diss.).

Leij, A. van der, Inprentingsmethodiek, informatieverwerking en differentiële instructiegevoeligheid, een aanzet tot theorievorming. In: L.M. Stevens, & A. van der Leij, (Red.), *Dyslexie 86.* Verslag van het congres te Utrecht. Lisse: Swets & Zeitlinger, 1986.

Leij, A. van der, & E.J. Kappers, Instructie van kinderen met ernstige leerproblemen: introductie op de thema-serie. *Pedagogische Studiën,* 1989, *66,* 311.

Leij, A. van der, & V.H.P. van Daal, Attacking dyslexia: the effect of proceding and simultaneous verbal practice. In: J J. Dumont & H. Nakken (Eds.), *Learning disabilities.* Vol. 2: *cognitive, social and remedial aspects.* Lisse: Swets & Zeitlinger, 1989, 179-191.

Perfetti, C.A., *Reading ability.* New York: Oxford University Press, 1985.

Perfetti, C.A., The presentation problem in reading acquisition. To appear in: Ph.B. Gough (Ed.), *Reading acquisition.* New York: Erlbaum (in press).

Reitsma, P., *Phonemic and graphemic codes in learning to read.* Amsterdam: Vrije Universiteit, 1983 (diss.).

Reitsma, P., Automatische behandeling van lees- en spellingsproblemen? In: A. van der Leij & L.M. Stevens (Red.), *Dyslexie.* Verslag van het congres te Den Haag. Lisse: Swets en Zeitlinger, 1985.

Sas, J.F. & C.J. Wieringa, DVO. *Het deelvaardigheidsonderzoek ten behoeve van het voorbereidend en aanvankelijk lees- en spellingsonderwijs.* Groningen: Wolters-Noordhoff, 1982.

Seymour, P.H.K., *Cognitive analysis of dyslexia.* Londen: Routledge & Kegan Paul, 1986.

Span, B., *Een informatieverwerkingsbenadering van cognitief functioneren. Deel 2: leren lezen.* Haren: RION, 1982.

Spear, L.C., & R.J. Sternberg, An information-processing framework for understanding reading disability. In: S.J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1987.

Spring, C. & C. Capps, Encoding speed, rehearsal, and probed recall of dyslexic boys. *Journal of Educational Psychology,* 1974, *66,* 780-786.

Spring, C. & L. Perry, Naming speed and serial recall in poor and adequate readers. *Contemporary Educational Psychology,* 1983, *8,* 141-145.

Struiksma, A.J.C., A. van der Leij & J. Vieyra, *Diagnostiek van technische lezen en aanvankelijk spellen.* Amsterdam: VU-boekhandel, 1987[2].

Torgesen, J.K., Computer-assisted instruction with learning-disabled children. In: J.K. Torgesen & B.L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities* New York: Academic Press, 1986, 417-435.

Curricula vitae

A. van der Leij (1946) studeerde psychologie aan de Universiteit van Amsterdam en werkte van 1971 tot 1977 bij het Gemeentelijk Schoolpsychologisch Bureau te Zwolle. Van 1977 tot heden is hij verbonden aan de Sectie Speciale Pedagogiek van de Vrije Universiteit, sinds 1984 als hoogleraar. Promoveerde in 1983 op een proefschrift onder de titel Ernstige leesproblemen.

H. Smeets (1956) studeerde orthopedagogiek met bijvakken sociaal-wetenschappelijke informatica en wijsgerig-empirische pedagogiek aan de RU te Leiden. Hij participeerde tijdens de studie in het door NWO gesubsidieerde onderzoek 'De kwaliteit van de hechtingsrelatie aan moeders, vaders en crêcheleidsters'. Vanaf 1988 werkt hij als wetenschappelijk onderzoeker bij de Vrije Universiteit, eerst met als opdracht de statistische dataverwerking van de Computergestuurde Toets voor Automatisering van het Lezen. In 1989 is hij verbonden aan het NWO-project 'Directe woordherkenning bij kinderen met leesproblemen'.

V. van Daal (1953) studeerde psychologische functieleer met bijvak onderwijskunde aan de Universiteit van Amsterdam. Verrichtte enige tijd onderzoek naar het bestuderen van studieteksten aan de Universiteit Twente. Vanaf 1985 werkt hij als onderzoeker aan de Vrije Universiteit (Sectie Speciale Pedagogiek) in projecten m.b.t. de ontwikkeling en beproeving van het Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen en de Computergestuurde Toets voor Automatisering van het Lezen. Vanaf 1988 is hij verbonden aan het NWO-project 'De Ontwikkeling van het Orthografisch Geheugen'.

Adres: Vrije Universiteit, Sectie Speciale Pedagogiek, Van der Boechorststraat 1, 1081 BT Amsterdam

Manuscript aanvaard 14-11-'89

Summary

Leij, A. van der, H. Smeets & V.H.P. van Daal. 'The development of the computer-assisted test for word recognition.' *Pedagogische Studiën,* 1990, *67,* 163-178.

The development of the computer-assisted test for word recognition (COTAL) is an example of a research project on the diagnosis of children with severe learning disabilities within a conceptual framework of information processing.

To indicate the specific characteristics of the reading process of children with severe reading disabilities, a computer-assisted test has been developed. Seven subtests reflect the various ways to present words as stimulus (auditory; visual) and the ways to respond (matching; naming; typing). Six wordcategories (one, two syllables) were used. In the reported study, 37 children with severe reading disabilities were compared to 37 reading age controls.

The results indicate that the tests reveal significant differences between very poor and normal readers, related to the use of the articulatory code. Implications for further research are discussed.

Leren denken binnen het vakgebied bestuurskunde: het leren oplossen van bestuurskundige ontwerpproblemen

C. TERLOUW
Onderwijskundig Centrum
Universiteit Twente, Enschede

Samenvatting

Op grond van een miniatuurtheorie voor onderwijsleerprocessen werden twee cursussen ontwikkeld voor het leren oplossen van bestuurskundige ontwerpproblemen. In deze cursussen staat het gebruik van een Systematische ProbleemAanpak (SPA) bij het leren oplossen van problemen centraal. Er werd voortgebouwd op ervaringen in natuurwetenschappelijke cursussen met een dergelijke SPA. De cursussen werden met vier groepen studenten in een veldexperiment (quasi-experimenteel design) summatief geëvalueerd. In drie van de vier groepen werden significante leereffecten gevonden aangaande het aantal systematische denkstappen. Er werd deels voldaan aan criteria ten aanzien van aanwezigheid en volgorde van bepaalde systematische denkhandelingen. Wij concluderen dat de cursussen effectief zijn en dat het mogelijk is het SPA-concept te generaliseren van natuurwetenschappelijk naar sociaalwetenschappelijk leren probleemoplossen.

Inleiding

Studenten in het hoger onderwijs stuiten op nogal wat moeilijkheden als zij problemen moeten oplossen in hun vakgebied. Tentamens die voornamelijk uit 'problemen' bestaan, zijn dan ook erkende struikelblokken voor studenten blijkens de lage slaagpercentages.

Onderzoek in het kader van het verbeteren van '(leren) probleemoplossen' kan verschillende vormen aannemen: (a) constaterend of construerend onderzoek; en (b) tekorten in de declaratieve kennis of in de procedurele kennis als centrale verklarende hypothese. Constaterend onderzoek beperkt zich tot de beschrijving van probleemoplossingsgedrag van studenten al dan niet in vergelijking tot expertgedrag. Construerend onderzoek gaat na of het leren van gewenste denkhandelingen voor het probleemoplossen tot betere prestaties leidt. In het geval dat de tekorten in de declaratieve kennis als primair worden gezien, zoekt men vooral de oplossing in een verbetering van het kennisverwervings- en verwerkingsproces. In het geval dat de tekorten in de procedurele kennis als primair worden gezien, is de centrale hypothese dat het probleemoplossen beter verloopt als wordt gewerkt met een systematische probleemaanpak. Een verschil van mening is voorts in welke mate een dergelijke systematische probleemaanpak domeingeoriënteerd moet zijn.

In het onderhavige onderzoek wordt uitgegaan van een construerend standpunt en de probleemaanpak-hypothese, zij het, dat in relatie daarmee, óók aan verwerking van de declaratieve kennis aandacht wordt besteed. De probleemaanpak in onze opvatting betreft een integratie van domeinspecifieke en niet-domeinspecifieke probleemoplossingsmethoden. Een construerend onderzoek impliceert, dat wordt uitgegaan van onderwijsontwikkeling. Wij gaan daartoe uit van de FUNDES-procedure (Terlouw, 1987). FUNDES is een samentrekking van FUNctional DESign, een systematische procedure voor onderwijsontwikkeling op handelingstheoretische grondslag (Terlouw, 1989), met als fasen probleemanalyse, ontwerp, constructie, uitvoering/formatieve evaluatie en beoordeling/summatieve evaluatie.

Uitgaande van een eerdere versie van de FUNDES-procedure en van de probleemaanpak-hypothese (met aandacht voor verwerking van declaratieve kennis) ontwikkelden Mettes en Pilot (1980) en Van Weeren, Kramer-Pals, De Mul, Peeters en Roossink (1982) effectieve natuurwetenschappelijke cursussen. Het meeste onderzoek beperkt zich tot het natuurwetenschappelijk domein (Vanderlocht & Van Damme, 1989).

In dit artikel rapporteren wij of een probleemaanpak-benadering ook effectief is in het sociaalwetenschappelijke domein. Het betreft in het bijzonder een onderzoek in de faculteit Bestuurskunde van de Universiteit Twente waarin met behulp van de FUNDES-procedure cursussen beleidsontwerpen werden ontwikkeld. Concreet betrof het twee cursussen van 125 uur voor 2de-jaars en 3de/4de-jaars studenten: ontwerpen van stadsvernieuwingsbeleid en ontwerpen van EG-beleid.

Achtereenvolgens komen aan de orde:

1. Doelen en vraagstellingen van het onderzoek;
2. Het theoretisch kader;
3. De onderzoeksopzet. Deze omvat de (on)afhankelijke variabelen, het design, de deelnemers en de methoden van dataverzameling;
4. De resultaten en conclusies; en
5. Discussie.

1 *Doelen en vraagstellingen van het onderzoek*

De *doelen* zijn praktisch en theoretisch van aard. De praktische doelstelling betreft de ontwikkeling van twee effectieve cursussen probleemoplossen in de faculteit Bestuurskunde van de universiteit Twente, in het bijzonder om studenten te leren beleid systematisch te ontwerpen.

De theoretische doelstelling is gericht op generalisatie van een instructie-opzet in het hoger onderwijs die is gebaseerd op het gebruik van een Systematische ProbleemAanpak (SPA) in het leren probleemoplossen.

De *vraagstellingen* zijn drieërlei:
Leidt een experimentele cursus waarin het gebruik van een Systematische ProbleemAanpak voor het beleidsontwerpen centraal staat bij studenten Bestuurskunde tot

a. een significante vergroting van het aantal systematische denkstappen bij het oplossen van een geoefend en een nieuw bestuurskundig ontwerpprobleem?
b. de aanwezigheid van bepaalde, gewenst geachte systematische denkstappen in de probleemaanpak?
c. de aanwezigheid van een bepaalde, gewenste volgorde van systematische denkstappen in de probleemaanpak?

2 *Theoretisch kader*

Het onderhavige onderzoek betreft een deelonderzoek binnen een groter onderzoek naar de bruikbaarheid van een handelingstheoretische onderwijsontwikkelingsprocedure voor het leren probleemoplossen in het hoger onderwijs, de hiervoor genoemde FUNDES-procedure (Terlouw, 1987). Binnen dit handelingstheoretische kader gaat het in deze rapportage om de bruikbaarheid van het onderwijsontwerp (fase 2).

Om tot een onderwijsontwerp te komen, zijn de volgende ontwerphandelingen relevant:

- de formulering van onderwijsleertheoretische uitgangspunten (2.1);
- de formulering van de doelstelling in de vorm van een Gewenst HandelingsVerloop (GHV) (2.2);
- de bepaling van de declaratieve en procedurele vakinhouden (2.3 en 2.4);
- de bepaling van het uit te voeren onderwijsleerproces (2.5).

2.1 *De onderwijsleertheoretische uitgangspunten*

Onderzoek laat zien dat de kennisbasis van experts op verschillende aspecten verschilt van die van novices en beginnelingen. Novices zijn studenten die net de cursus hebben afgesloten, terwijl beginnelingen net met de cursus zijn begonnen. In het onderhavige geval zijn wij vooral geïnteresseerd in het gedrag van beginnelingen. Het ontbreekt beginnelingen aan met elkaar verbonden schemata van declaratieve en procedurele kennis. Beginnelingen werken voorts langzamer, maken meer fouten en gebruiken een weinig systematische benadering (Chi, Glaser & Rees, 1982); Larkin, 1980). Uitgaande van een handelingstheoretische benadering (Van Parreren, 1981, 1982, 1988; Van Parreren & Schouten-Van Parreren, 1981) gaan wij uit van drie onderwijsleertheoretische uitgangspunten om beginnelingen probleemoplossen op hun vakgebied te leren.

In de *eerste* plaats moet de student een goede oriënteringsbasis opbouwen. Een student beschikt over een goede oriënteringsbasis voor een uit te voeren handeling, als hij zo volledig mogelijk weet wat en hoe hij het moet doen en daarvan de zin ziet (Van Parreren, 1988).

De oriënteringsbasis voor het beleidsontwerpen werd gevormd door de procedurele kennis (de probleemaanpak) met daaraan gekoppeld de declaratieve kennis. Het laatste type kennis werd geordend aan de hand van de fasen van de probleemaanpak. De oriënteringsbasis voor een beleidsontwerpprobleem is met opzet onvolledig (Tichomirov, 1974; Podd'jakov, 1979; Nelissen, 1980). Het behoort tot de aard van een beleidsontwerpprobleem dat het te bereiken doel en de weg waarlangs dit is te bereiken, tevoren niet precies vast staan. Veelal moet daarover in een sociaal proces eerst nog overeenstemming worden bereikt. Bovendien bevat het bestuurskundige domein (volgens sommige bestuurskundigen in principe) maar weinig generaliseerbare kennis over middel-doel relaties. Een principiële openheid is zeker nodig om een eigen zingeving voor het te bereiken doel en de gekozen weg op te bouwen. Door deze onvolledige oriënteringsbasis moet na afloop van het onderwijs tenminste een oriënteringachteraf (reflectie) plaatsvinden op de procedurele en declaratieve kennis ten behoeve van het te bereiken doel, de gekozen weg en de zingeving daarvan. In deze reflectie wordt op basis van de voorgaande ervaringen de vooraf gegeven oriënteringsbasis verbeterd (Zak alsmede Jantos in Nelissen & Vuurmans, 1983; Nelissen, 1987).

In de *tweede* plaats moet de student met de opgebouwde oriënteringsbasis gefaseerd oefenen. De fasen hebben betrekking op het abstractieniveau waarop de handelingen uit de oriënteringsbasis worden geoefend (bijv. het mentale, verbale en gematerialiseerde niveau van handelen). Voor het beleidsontwerpen lieten we studenten op basis van de cyclus van ervaringsleren (Erkamp, 1980) zowel zelfstandig als in taakgroepen oefenen met de aangereikte oriënteringsbasis om tot een optimaal coöperatief probleemoplosproces te komen (Jantos, 1982). De fasering betrof een afwisseling van de genoemde abstractieniveaus (Kossakowski, 1982).

In de *derde* plaats moet er per leerfase worden getoetst en op systematische wijze terugkoppeling worden gegeven. Een leerfase heeft betrekking op een vakinhoudelijk afgerond geheel waarop is georiënteerd en waarmee is geoefend. De toetsing en terugkoppeling sluit de leerfase af en gaat na of kan worden doorgegaan naar de volgende vakinhoudelijke eenheid. Voor het beleidsontwerpen werden per leerfase de produkten van studenten getoetst en terugkoppeling gegeven aan de hand van de aangereikte probleemaanpak. Er werd geen strikte doorgaan/niet-doorgaan regel gehanteerd; men liet studenten 'reparaties' van produkten zoveel mogelijk in de volgende leerfase integreren.

2.2 *Het Gewenst HandelingsVerloop*

In paragraaf 2.1 is gesteld dat declaratieve en procedurele vakinhouden gerelateerd moeten worden aangeboden. De formulering van een Gewenst HandelingsVerloop geeft een kader voor een dergelijke relatering. Een Gewenst HandelingsVerloop is een doelstellingenformulering waarin met psychologische termen de handelingen die verricht moeten worden om de gewenste leerresultaten te bereiken, op een samenhangende wijze staan aangegeven (Terlouw, 1987).

Wij volgden de methode van de rationele reconstructie (Resnick, 1976) en literatuurstudie voor de formulering van het Gewenst HandelingsVerloop en valideerden het resultaat met behulp van een onderzoek naar het handelingsverloop bij experts en bij beginnelingen (Terlouw & Mettes, 1986). Voor de precieze werkwijze en het daaruit voortkomende stroomschema zij verwezen naar Terlouw, Mettes en Roemers (1981) en Terlouw (1987). De docenten accepteerden de doelstelling in de vorm van een Gewenst HandelingsVerloop.

2.3 *De declaratieve vakinhouden*

Het gaat hier om het beschrijvende deel van het kennisrepertoire dat de student nodig heeft om een probleem op te lossen. Het betreft feiten, regels, definities van concepten, formules, theorieën, e.d. Gelet op de hiervoor genoemde openheid van het probleem was het niet mogelijk een afgerond geheel van essentiële kenniselementen te formuleren. Daarom werd uitgegaan van een programma van startkennis. Dit is een geheel van kennis dat het mogelijk maakt in ieder geval met de aanpak van een probleem een start te maken. In de loop van het onderwijs konden studenten zelf hun kennisbasis op eigen wijze uitbreiden met de startkennis als springplank. In het onderhavige geval bestond het programma van startkennis uit een lijst met specifieke kennisthema's die moesten worden bestudeerd. Der-

gelijke thema's werden in een syllabus aan de hand van het Gewenst HandelingsVerloop geordend aangeboden. De kennisverwerving werd getoetst. Studenten konden alléén deelnemen aan het onderwijs als een voldoende resultaat op deze toets werd gehaald.

2.4 *De procedurele vakinhouden: de Systematische ProbleemAanpak*

Het concreet oplossen van een probleem vraagt om een hanteerbaar *leermiddel* dat studenten als hulp kunnen gebruiken om de kans op een adequaat handelingsverloop zo groot mogelijk te maken. Daarom werd een dergelijk leermiddel uit het GHV afgeleid. Wij noemen dit leermiddel een *Systematische ProbleemAanpak (SPA)*. Een SPA is een (leer)middel ter oriëntering van de student voor het probleemoplossen. In het onderhavige geval van beleidsontwerpen bestaat de SPA uit een systeem van algemene en domeinspecifieke probleemanalyse-methoden. Gegeven de openheid van het probleem kunnen slechts analyse-aanwijzingen worden gegeven om doel en weg in kaart te brengen en daarin keuzen te maken. De gehele SPA kan worden gezien als een heuristiek om een slechtgedefinieerd probleem te transformeren in een goed-gedefinieerd probleem. Wij beschouwen deze SPA als de procedurele vakinhoud. Na enige proefversies (zie Terlouw, 1982; Terlouw, 1987) werd de definitieve vormgeving gevonden in de vorm van een checklist met vaste hoofdpunten en mogelijke aandachtspunten. Een beschrijving:

1. *Bepaal de aanwijzingen uit de ontwerpopdracht*

Het gaat hier om het analyseren van de ontwerpopdracht en andere relevante documenten. Mogelijke categorieën zijn: problemen, doeleinden, middelen, etc. Er kan in algemene zin, maar ook vanuit een rol geanalyseerd worden. In het geval van onduidelijkheden neme men contact op met de opdrachtgever.

2. *Globaal overzicht van ontwerpprobleem en situatie*

In het eerste deel gaat het om een grondige probleemanalyse in termen van deelproblemen en een bestaande en gewenste situatie. In het tweede deel gaat het om een situatieanalyse: een analyse van het krachtenveld waarin het probleem ligt ingebed. 'Beleidsvrijheid', 'machtsverdeling', 'belangen', e.d. zijn mogelijke dimensies bij de analyse van het krachtenveld.

3. *Een hypothetisch causaal model opstellen*

Centraal staat het bepalen van de mogelijke oorzaken van het probleem op basis van de literatuur en naar eigen inzichten. Met behulp van de aangegeven punten kan een eigen (bescheiden) causaal model van hypothetische aard worden opgezet.

4. *Een beschrijving van de einddoelen*

Het gaat hier om een zo empirisch mogelijke beschrijving van de einddoelen. Einddoelen kunnen per actor verschillen. Voorts hangen einddoelen samen met gekozen maatstaven, het opgestelde causale model en de geldende randvoorwaarden.

5. *Planning van de communicatie*

Het betreft zowel de mondelinge als de schriftelijke communicatie. De aandachtspunten zijn handreikingen om de communicatie voor te bereiden. 'Vaste structuren' betreffen vaste communicatiepatronen, te vergelijken met een script (Steehouder, Jansen, Van der Staak & Woudstra, 1984). Voorts is er in dit kader speciale aandacht voor het onderhandelen.

6. *Evaluatie*

Het gaat hier om de evaluatie van zowel allerlei (tussen)produkten als van het verloop van ontwerpprocessen zelf. Bij evaluatie spelen criteria een belangrijke rol.

7. *Bestaande beleidsmiddelen gebruiken*

Er moet kritisch worden nagegaan of bestaande middelen zijn te gebruiken. Geldigheid en kosten/baten zijn in dit kader belangrijke overwegingen. Een model van beleidseffecten – een explicitering van de intermediërende processen die tot het gewenste effect leiden (Hoogerwerf, 1984) – kan behulpzaam zijn bij de keuze van een beleidsmiddel.

8. *Nieuwe beleidsmiddelen met behulp van beleidsfuncties bedenken*

Beleidsfuncties zijn algemene beschrijvingen van wat er beleidsmatig moet gebeuren om van een bestaande situatie in een gewenste situatie te geraken. Dergelijke algemene transformaties dienen als 'trigger'-mechanisme om in een brainstorm-sessie of via de literatuur

concrete, alternatieve, nieuwe beleidsmiddelen te vinden (Terlouw, e.a., 1981). Voor de rest is de werkwijze zoals beschreven bij 8.

9. *De voorgenomen plannen uitvoeren en integreren*

Eerder zijn allerlei plannen opgesteld: voor probleemanalyse; voor communicatie, voor evaluatie, e.d. Deze worden nu op elkaar afgestemd en daadwerkelijk uitgevoerd. De eisen van de te ontwerpen beleidsnota spelen hierin een belangrijke rol. Voorts moet er aandacht zijn voor de gevolgen van een bepaalde deeluitkomst voor het totaalbeeld dat men van de situatie heeft.

10. *De laatste controles en wat men ervan heeft geleerd*

Het gaat uiteindelijk om het uitbrengen van een beleidsnota. Voordat deze uitgaat moet deze worden gecontroleerd op volledigheid, interpunctie, spelling, e.d. Tot slot wordt in de tijd teruggekeken en gereflecteerd over de processen die goed en fout liepen en wat dat betekent voor een in de toekomst te volgen probleemaanpak.

Voor wat betreft het gebruik van de SPA kregen de studenten een aantal adviezen over de handelingsvolgorde, het cyclische karakter van het handelen en de koppeling van handelingen.

2.5 *Het model van onderwijsleerprocessen*

Gelet op het voorafgaande werd een miniatuurtheorie (De Klerk, 1983) geformuleerd voor de gewenste onderwijsleersituatie. Wij noemen dit het model van onderwijsleerprocessen. Dit model fungeerde als leidraad voor de verdere cursusontwikkeling en het didactisch handelen. Het model van onderwijsleerprocessen staat in Figuur 1. Voor de verschillende termen in Figuur 1 zij verwezen naar de paragrafen hiervoor.

3 *Onderzoeksopzet*

De cursusopzet in paragraaf 2 was het resultaat van vier jaren onderwijsontwerp, onderwijsconstructie, onderwijsuitvoering en formatieve evaluatie (de FUNDES-fasen 2, 3 en 4). Deze onderwijsontwikkeling vond steeds in nauwe samenwerking met de onderwijspraktijk plaats. Voor een onderbouwing van de claim dat het onderwijs effectief is, moet men kunnen beargumenteren dat betere leerprestaties zijn toe te schrijven aan het gegeven onderwijs. Andere verklaringsbronnen moet men kunnen uitsluiten. Dit nu is het doel van de vijfde FUNDES-fase, de summatieve evaluatie.

Achtereenvolgens zullen wij op de volgende aspecten ingaan: de onafhankelijke en afhankelijke variabelen en de criteria (3.1); het onderzoeksdesign (3.2); de deelnemers (3.3); en de methoden van dataverzameling (3.4).

3.1 *De onafhankelijke en variabelen en de criteria*

De *onafhankelijke variabele* betrof twee op basis van de miniatuurtheorie ontwikkelde probleemoploscursussen 'Bestuurskundig ontwerpen' voor 2e en 3e/4e-jaars studenten Bestuurskunde aan de universiteit Twente. Eén cursus richtte zich op het ontwerpen van stadsvernieuwingsbeleid; de ander op het ontwerpen van EG-landbouwbeleid. Voor de gedetailleerde onderwijsplannen zij verwezen naar Terlouw (1987). Ze zullen nu summatief worden geëvalueerd met *nieuwe groepen* studenten bestuurskunde.

De *afhankelijke variabele* was, gelet op onderzoeksvraagstellingen in paragraaf 1, drieërlei:

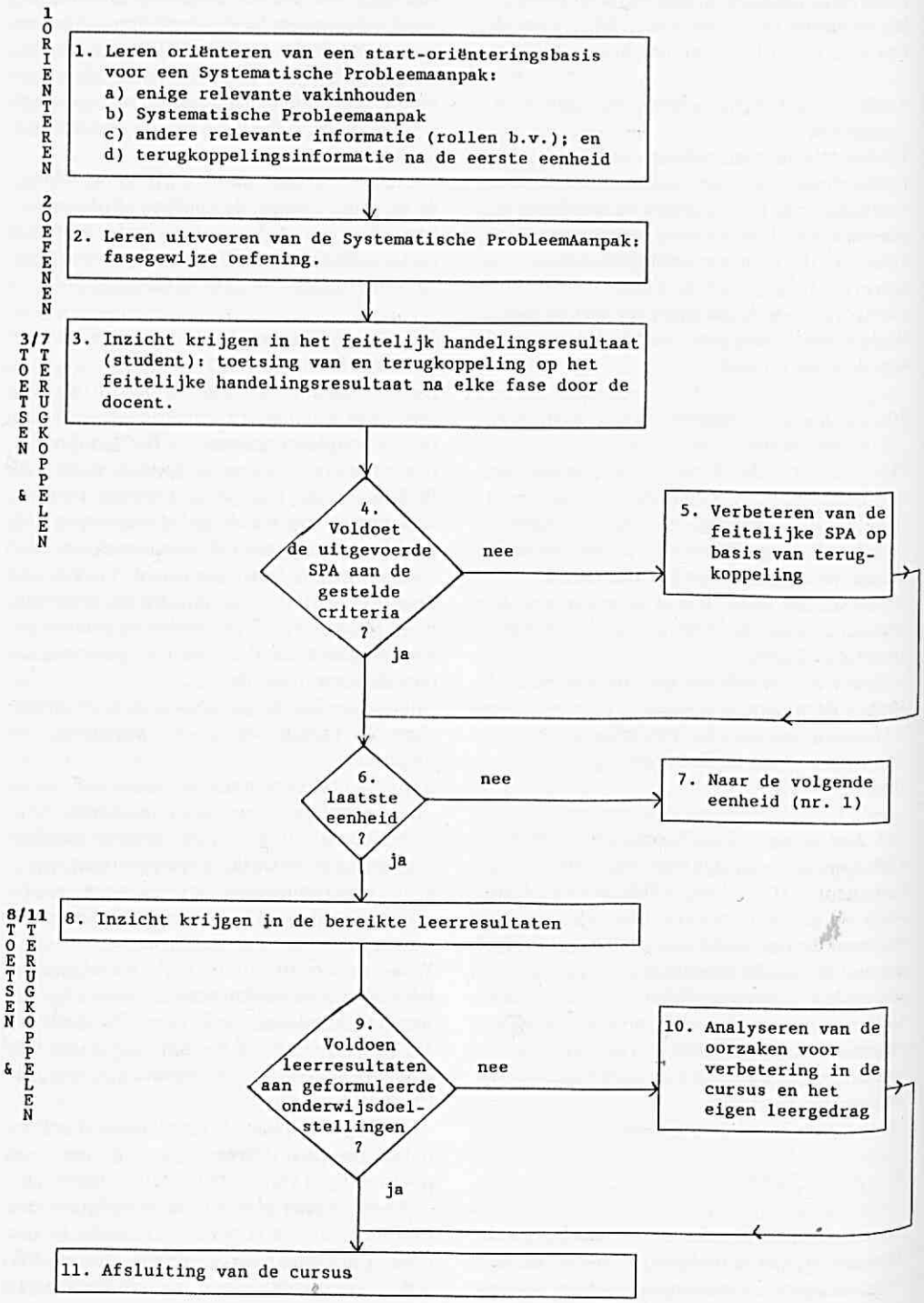
a. het aantal systematische denkstapppen in een geoefend en een nieuw probleem;
b. het aantal studenten dat gewenst geachte systematische denkstappen gebruikt; en
c. het aantal studenten dat een gewenst geachte volgorde in de systematische denkstappen gebruikt.

Waarom deze drie afhankelijke variabelen op het niveau van denkprocessen (beter: het inwendig handelingsplan; zie Ponomarev, 1973) en bijvoorbeeld *niet* het produkt dat uiteindelijk het resultaat is van het onderwijs, nl. een beleidsnota?

In de eerste plaats is de gekozen operationalisering in overeenstemming met de doelstellingen van het betreffende onderwijs.

In de tweede plaats is de beleidsnota een produkt van een complexe interactie in een open simulatie waarop slechts gedeeltelijk werd voorbereid door de voorafgaande analyse.

Bovendien was er in de derde plaats een vakinhoudelijk beoordelingsprobleem. De beleidsnota kon eigenlijk alleen maar goed be-



Figuur 1 *Model van onderwijsleerprocessen: een miniatuurtheorie voor de onderwijsleersituatie van 2e t/m 4e-jaars-studenten Bestuurskunde in een cursus Bestuurskundig probleemoplossen*

oordeeld worden in het licht van de in de simulatie gegroeide politieke/bestuurlijke situatie. Een technisch slechte nota kon politiek gezien uitstekend zijn en omgekeerd. Deze beoordelingssituatie was te complex om tot valide en betrouwbare oordelen van docenten te komen.

Op grond van deze drie overwegingen beperkten wij ons tot de drie genoemde variabelen op het niveau van de denkprocessen. Voor elk van de afhankelijke variabelen werden tevoren criteria vastgesteld:

ad a) De verschillen tussen voor- en nametingen moeten tenminste significant zijn op 5%-niveau;

ad b) en c) Op basis van de Systematische ProbleemAanpak (zie par. 2.4) werd een scoringsvoorschrift geformuleerd waarin stond gespecificeerd welke SPA-stappen in welke volgorde moesten voorkomen. Vervolgens werden normen geformuleerd.

Een voorbeeld: Na de cursus moet 70% van de studenten van de volgende zes handelingen er tenminste vier gebruiken, met dien verstande, dat van deze vier de eerste drie altijd moeten worden gebruikt: 1) aanwijzingen bepalen in de opdracht; 2) analyseren van een gekozen deelprobleem; 3) analyseren van het krachtenveld; 4) opstellen van een hypothetisch causaal model; 5) einddoelen formuleren; 6) planning van de communicatie (zie voor een volledig overzicht van de criteria Terlouw, 1987).

3.2 *Het onderzoeksdesign*

Er is sprake van een summatieve evaluatie. De onderzoeksopzet betrof een quasi-experimenteel cohort design: een combinatie van een 'recurrent institutional cycle design' en een 'institutional cycle design' die elkaar wederzijds versterken aangaande interne en externe validiteit (Cook & Campbell, 1979). Het totale design is in wezen een combinatie van een post-test only design with nonequivalent groups en een one-group pretest-posttest design. Het design staat weergegeven in Figuur 2.

Er werden vier soorten van metingen verricht: een voortoets, een natoets, een transfertoets en een uitgestelde transfertoets. Afhankelijk van het cohort waarin de student zich bevond, werd een bepaalde combinatie van deze toetsen afgenomen. De metingen waren zodanig gepland dat de interne, externe en construct validiteit na afloop van het onderzoek op eventuele storingsbronnen konden worden gecontroleerd. Enige voorbeelden (zie Figuur 2):

- History, testing, instrumentation, de interaction of selection and maturation, de interaction of testing X en de reactive arrangements zijn bij de groepen niet-1ste-fase studenten gecontroleerd als de vergelijking $O_{84/85}\ EG_{begin} < O_{83/84}\ EG_{eind}$ opgaat. Als de vergelijkingen $O_{83/84}\ SVN_{begin} < O_{82/83}$ Info, $O_{83/84}\ EG_{begin} < O_{82/83}$ Info, $O_{84/85}\ EG_{begin} < O_{83/84}$ Info opgaan, is dat een ondersteuning voor de claim dat deze storingsbronnen zijn gecontroleerd.
- Selection wordt niet gecontroleerd door bovengenoemde vergelijkingen. Maar dit is gecontroleerd als de vergelijkingen $O_{83/84}\ SVN_{begin} < O_{83/84}\ SVN_{eind}$ én $O_{83/84}\ EG_{begin} < O_{83/84}\ EG_{eind}$ óók bewaarheid worden.

Campbell en Stanley (1966) en Cook en Campbell (1979) geven schema's voor dergelijke vergelijkingen. Zowel voor niet-1ste-fase studenten als voor 1ste-fase studenten zijn dergelijke schema's geformuleerd.

Voorts werden er ter controle van de mate en de kwaliteit van de implementatie van de experimentele cursussen, de onafhankelijke variabele, eveneens metingen gepland.

3.3 *De deelnemers*

Er was sprake van twee cursussen die aan drie soorten van groepen in de faculteit Bestuurskunde van de Universiteit Twente werden gegeven (N = populatie-aantal; n = steekproef-aantal; het subscript 1 of 2 geeft aan of er respectievelijk sprake is van een groep die alle toetsen ondergaat of van een groep die, in verband met de controle op instrumentatie, alléén de transfertoets ondergaat):

- een formatieve groep 3e/4e-jaars studenten in het studiejaar 1982/83. Dit is nog een oude versie van de twee cursussen ($N = 63$; $n = 16$);
- een experimentele groep 3e/4e-jaars studenten "oude stijl" in het studiejaar 1983/84 (het eerste en het tweede trimester). Het grootste deel hiervan kon uit de cursussen een keuze maken ($N_{SVN} = 36$; $n_1 = 13$; $n_2 = 8$) ($N_{EG} = 34$; $n_1 = 15$; $n_2 = 8$); een beperkt aantal studenten volgde beide cursussen ($N = 21$; $n_1 = 5$; $n_2 = 8$);
- een experimentele groep 2e-jaarsstudenten "eerste fase" in het studiejaar 1983/84.

Studiejaar: / Jaren: / trimester: / Cohort deelnemers generatiejaar:	'82 1	'82/'83 2	'82/'83 3	'83 1	'83/'84 2	'84 3	'84/'85 1
A. niet 1e Fase studenten							
1979/1980 (of ouder)	X_1		Y_1 O_{info}	O_{svn} X_2 O_{svn} O_{info}		O_{info} (uitgesteld) SVN-groep	
1980/1981 (of ouder) in 2 groepen					O_{eg} Y_2 O_{eg} O_{info}	O_{info} (uitgesteld) EG-groep	
							O_{eg}
1981/1982			O_{svn} O_{info}				
			O_{eg} O_{info}				
B. 1e Fase studenten 1982/1983						O_{svn} X_2 O_{svn} O_{info}	O_{info} (uitgesteld) SVN-groep
						O_{eg} Y_2 O_{eg} O_{info}	O_{info} (uitgesteld) EG-groep

Toelichting:
X_1 = 1982/1983 cursus Stadsvernieuwing
Y_1 = 1982/1983 cursus EG-Beleid
X_2 = 1983/1984 cursus Stadsvernieuwing
Y_2 = 1983/1984 cursus EG-Beleid
O_{svn} = voor/natoets cursus Stadsvernieuwing
O_{eg} = voor/natoets cursus EG-Beleid
O_{info} = transfertoets (Informatie-casus)
N.B.: voor alle O's zijn deelnemers random uit de groep getrokken.

Figuur 2 *Overzicht van het totale design van de summatieve evaluatie*

Deze groep kon uit twee cursussen een keuze maken ($N_{SVN} = 59$; $n_1 = 10$; $n_2 = 11$) ($N_{EG} = 72$; $n_1 = 9$; $n_2 = 10$).
Ten slotte was er nog een controlegroep 2de-jaars studenten die in dat jaar geen cursus volgde, omdat zij nog tot de niet-eerste-fase studenten behoorden ($N = 120$ potentieel; $n_1 = 6$; $n_2 = 6$). Het totale aantal getoetste studenten betrof 143.

Voor de totale groep was er dus geen sprake van random toewijzing. Voor de geplande metingen werden uit deze groepen wel op random wijze steekproeven getrokken. De studiesnelheid werd als een mogelijke storingsbron gezien. Bijgevolg werden daarover gegevens verzameld om deze te gebruiken als covariaat in de analyse.

3.5 *Methoden van dataverzameling*
Er was sprake van twee evaluatieplannen: één voor het verzamelen van gegevens over de mate en kwaliteit van de implementatie van de onderwijsleerprocessen (het onderwijsplan), de onafhankelijke variabele en één voor het verzamelen van de leerresultaten, de afhankelijke variabele.
ad 1. *Evaluatieplan implementatie onderwijsleerprocessen*
Een evaluatie-draaiboek en een vragenlijst werden opgesteld op basis van het model van onderwijsleerprocessen (zie Figuur 1, par. 2.5) en het daarbij horende onderwijsplan. Tijdens de onderwijsuitvoering beoordeelden twee onafhankelijke observatoren op een driepuntsschaal in een evaluatiedraaiboek, òf een gepland onderwijsmiddel werd uitgevoerd en zo dit het geval was, de kwaliteit van de onderwijsfunctievervulling. Direct na afloop van het onderwijs vulden studenten een vragenlijst in, waarin hun oordeel werd gevraagd over de kwaliteit van de uitvoering en de functionaliteit van de geplande onderwijsmiddelen en -procedures. In meerderheid betroffen dit gesloten vragen met een vijfpuntsschaal.
ad 2. *Evaluatieplan leerresultaten*
Er werden twee instrumenten gebruikt: een combinatie van hardopdenken en interview en een vragenlijst. Een combinatie van hardopdenken en interview werd gebruikt om een zo compleet mogelijk beeld te krijgen van de probleemaanpak van de studenten. Voor het interview werd gebruik gemaakt van gestandaardiseerde vraagvormen (Laurillard, 1979). Voor de wijze van protocolanalyse volgden wij de methodologie van Breuker (1981). Ten behoeve van de protocolanalyse werd een interpretatiemodel geformuleerd op basis van de procedure vakinhouden (de SPA) (zie par. 2.4).

De problemen die werden voorgelegd betroffen drie casussen: twee geoefend, stadsvernieuwing en EG-Beleid; één geheel nieuw: Ontwerp een recht op informatie. De procedure van protocolafname was in een proefleider-instructie vastgelegd.

In de vragenlijst werden de leerresultaten op twee manieren gemeten: enkele gesloten vragen met vijfpuntsschalen en met behulp van een 'learner report' (Ik heb geleerd dat ...; zie De Groot, 1980).

Verschillende instrumenten werden ingezet om de resultaten wederzijds te valideren (Denzin, 1978; Cronbach, 1983).

4 *Resultaten en conclusies*

Eerst zullen we in drie paragrafen de uitkomsten van de drie gespecificeerde afhankelijke variabelen bespreken, respectievelijk, het *aantal denkstappen* (SPA-handelingen) (4.1), de *aanwezigheid* bij studenten van de *gewenste SPA-handelingen* (4.2) en de *aanwezigheid* bij studenten van de *gewenste volgorde in SPA-handelingen*. Ten slotte rapporteren we over de controle van de alternatieve verklaringen.

4.1 *Het aantal denkstappen (SPA-handelingen)*

De betrouwbaarheid van scoring werd nagegaan met de overeenstemmingsmaat K (Bishop, Fienberg & Holland, 1975). Er werd een K gevonden van .72 (40 proefpersonen; 843 protocoluitspraken). Ter bepaling van het effect van een probleemoploscursus op het aantal gebruikte systematische denkstappen, werden vier herhaalde metingen designs geconstrueerd (Winer, 1970) en vervolgens geanalyseerd met behulp van een covariantieanalyse uit de MANOVA-procedure (SPSS-X; Norušis, 1985). Er werd gebruik gemaakt van een herhaalde metingen design, omdat er sprake is van meer dan twee metingen per persoon. In een dergelijk geval is een herhaalde metingen design een betere strategie dan het toepassen van gecorreleerde T-toetsen (Norušis, 1985). Een overzicht van de resultaten per experimentele cursusgroep staat in Tabel 1.

Onze conclusies:
a. Drie van de vier experimentele cursussen hebben effect op het aantal systematische denkstappen ($p<.005$; zie in Tabel 1 onder 4c en 4d). De afwezigheid van een significant effect in de vierde cursus kon worden verklaard op basis van de verzamelde implementatiegegevens: een beperkte implementatie, in het bijzonder van juist die onderwijsmiddelen die vanuit de onderwijsleertheorie als van groot belang worden gezien (het oefenen).
b. De studiesnelheid (de covariaat) heeft geen invloed op de resultaten.
c. Met behulp van symmetrietests zijn de assumpties aangaande de contrasten nagegaan

Tabel 1 *Overzicht van uitkomsten van vier herhaalde metingen designs van vier cursusgroepen.*
*$p<0.5$; ** $p<.01$; *** $p<.005$

Bereken de waarden van REP éénweg-designs \ Cursusgroepen	Groep oude stijl 3e/4e jaar 1e trimester 1983/1984 Stadsvernieuwing	Groep oude stijl 3e/4e jaar 2e trimester 1983/1984 EG-Beleid	Groep 1e fase 2e jaar 3e trimester 1983/1984 Stadsvernieuwing	Groep 1e fase 2e jaar 3e trimester 1983/1984 EG-Beleid
1. n personen	13	15	9	8
2. gemiddeld aantal SPA-stappen (standaard-afwijking) per trimester/vakinhoud-groep				
a. casus vooraf	4.4 (1.4)	4.9 (1.5)	4.9 (1.2)	5.6 (2.1)
b. casus achteraf	8.0 (2.4)	7.1 (1.8)	7.2 (1.8)	7.1 (1.8)
c. transfercasus	6.1 (2.0)	7.3 (1.7)	6.6 (1.9)	7.4 (1.6)
3. gemiddelde (standaard-afwijking) constante covariaat "studie-snelheid"	4.9 (.9)	4.7 (.6)	4.5 (.9)	4.8 (.8)
4. F-waarden (overschrijdingskans) significantietesten				
a. constant	6.7 (.025)*	8.8 (.011)*	9.6 (0.17)*	2.0 (.202)
b. regressie covariaat	0.0 (.998)	.19 (.671)	.21 (.662)	.70 (.436)
c. multivariate tests casus (benaderde F)	10.3 (.003)***	19.2 (.000)***	26.6 (.001)***	1.91 (.228)
d. univariate contrast tests: Difference				
d1.(cas. vooraf) vs. (cas. achteraf) Helmert	21.9 (.001)***	14.7 (.002)***	49.0 (.000)***	3.5 (.104)
d2.(cas. vooraf) vs. cas. achteraf + transfercasus/2	20.3 (.001)***	34.9 (.000)***	12.0 (.009)***	4.4 (.073)
d3.(cas. achteraf) vs. (transfercasus)	7.3 (.019)*	.10 (.754)	.80 (.397)	.16 (.699)
e. averaged univariate tests	13.5 (.000)***	12.4 (.000)***	5.8 (.013)*	3.04 (.080)

(significantieniveau 5%). Slechts voor één groep (groep 1e fase, 2e jaar; 3e trimester 1983/84 Stadsvernieuwing) werd aan de assumpties niet voldaan. Dit betekent dat de univariate benadering in Tabel 1 onder e in drie gevallen een extra bevestiging vormen van de wel of niet gevonden effecten.

d. In geconstrueerde tweeweg- en drieweg-designs werd op exploratieve wijze de mogelijke invloed van de bestuurskundige vakinhoud (stadsvernieuwing of EG-beleid) en de jaargroep (2e of 3e/4e-jaars) nagegaan. Deze variabelen hadden geen invloed op de resultaten (voor een uitgebreide bespreking zie Terlouw, 1987).

4.2 *De aanwezigheid van gewenste SPA-handelingen*

De handelingen in de protocollen werden op twee niveaus geanalyseerd: macro- en micro-niveau. Het macro-niveau betrof de drie algemene delen van een handeling: oriënterend, uitvoerend, en controlerend deel. Het micro-niveau betrof de specifieke probleemaanpak-handelingen (SPA) (zie par. 2.4).

Kort samengevat zijn de conclusies als volgt:

a. Zowel in de geoefende casussen als in de nieuwe transfercasus zijn over alle cursussen gezien bij precies 70% (mediaan) van de studenten de drie gewenste handelingsdelen op macro-niveau – oriënteren, uitvoeren en controleren – in de probleemaanpak aanwezig. Per cursus zijn er evenwel verschillen. In vergelijkbare groepen die de cursus *niet* hebben gevolgd is bij slechts 33% (mediaan) van de studenten een dergelijke handelingsstructuur aanwezig. Op dit globale handelingsniveau is het gestelde evaluatie-criterium bereikt.

b. Op micro-handelingsniveau – de specifieke SPA-handelingen – zijn vooral de oriënterende SPA-handelingen aanwezig. Alleen in het geval dat cursusgroepen een nieuwe transfercasus als uitgestelde nameting ondergingen, werd het gestelde 70%-criterium gehaald. In deze groepen waren bij 73% (mediaan) van de studenten specifieke gewenste, oriënterende SPA-handelingen aanwezig. Opmerkelijk is voorts dat alleen in deze groepen relatief meer SPA-handelingen uit het uitvoerend deel worden gebruikt. Er is wel sprake van een stijging in de mediane percentages als we de voormeting met de nameting vergelijken van de geoefende casussen (van 15.7% naar 42%), echter, het 70%-criterium wordt niet gehaald. Wèl is het zo dat de formatieve cursusgroep en de niet-cursusgroepen achterbleven met mediane percentages van respectievelijk 25% en 27%.

4.2 *De aanwezigheid van een gewenste volgorde in SPA-handelingen*

Ook hier werden de protocollen op macro- en microniveau geanalyseerd.

De conclusies zijn als volgt:

a. Bij 70.3% (mediaan) van de studenten van de cursusgroepen is al in de voormeting een gewenste macro-volgorde – oriënteren vóór uitvoeren en controleren vrij – aanwezig. In de nameting van de cursusgroepen met de geoefende casus en de nieuwe transfercasus is sprake van enige stijging in de mediane percentages naar respectievelijk 75.4% en 72.9%, echter, deze stijging is zo beperkt dat we daar weinig waarde aan hechten. Wel is het zo dat het mediane percentage van de niet-cursusgroepen in de nieuwe transfercasus achterblijft en onder de 70%-norm blijft ten opzichte van de cursusgroepen (65% versus 72.9%). Ook hier zijn de verschillen echter weer klein.

b. Op micro SPA-handelingsniveau wordt het 70%-criterium zowel bij de voormeting als bij de nametingen door geen enkele groep gehaald. Wel is er weer in de cursusgroepen sprake van enige stijging in het mediane percentage van voormeting naar nameting, namelijk van 52% in de voormeting naar 60.7% in de geoefende casus en 56.6% in de nieuwe transfercasus. Ook hier blijven de niet-cursusgroepen qua mediaan percentage achter met respectievelijk 53.5% en 51.6%. Echter, de verschillen zijn klein en het criterium wordt niet gehaald.

4.4 *Controle op ondersteuning en alternatieve verklaringen*

Er is sprake van een quasi-experimentele opzet. Daarom is het noodzakelijk te controleren op alternatieve verklaringen voor de gevonden effecten. Daartoe wordt ingegaan op de implementatiewaarde van de cursussen, de mate van convergentie met andere metingen en de plausibiliteit van alternatieve verklaringen.

ad 1 *De implementatiewaarde van de cursussen*

In de voorgaande paragrafen zijn effecten geconstateerd. Om een causale relatie tussen het gegeven onderwijs en het gevonden effect te kunnen leggen, moet worden aangetoond dat het onderwijs volgens plan is uitgevoerd en van voldoende kwaliteit is. Wij noemen dit de implementatiewaarde van het onderwijs. De resultaten zijn als volgt:

Over het algemeen werden de cursussen uitgevoerd volgens plan. De uitval aan studenten was gering (n = 8) en had niet te maken met de experimentele behandeling. Alle vaste opdrachten werden uiteindelijk met een voldoende resultaat afgerond. Een kleine 10% van de geplande activiteiten werd niet uitge-

voerd; in enkele gevallen was dat functioneel. De betrouwbaarheid van het oordeel 'wel of niet uitgevoerd' was bevredigend: in 93% van de gevallen kwamen de twee observatoren tot hetzelfde oordeel. De kwaliteit van de cursusuitvoering was over het algemeen voldoende. Alleen bij de 2ejaars cursus EG-Beleid (in het derde trimester 1983/84) rees twijfel over de kwaliteit. In alle gevallen leek de schending van de kwaliteit echter gering en vond zij slechts plaats in de uitvoering van een bepaald onderwijsmiddel (b.v. de simulatie). Observatoren-data en studentenoordelen convergeerden ten aanzien van de uitvoeringskwaliteit ($r = -.61$).

Per saldo waren naar het oordeel van de studenten (vragenlijst) de onderwijsfuncties voor het grootste deel gerealiseerd: de gestelde evaluatie-criteria werden bereikt. De docenten waren het daarmee eens. Onze conclusie is dat de implementatiewaarde hoog genoeg is voor de claim dat er een causale relatie bestaat tussen het gegeven onderwijs en de gevonden effecten.

ad 2 *Convergentie met andere metingen*
De claim dat er effecten zijn, staat sterker als leereffecten óók met andere soorten instrumenten worden gevonden. Is dit niet het geval, dan moet men de effecten relativeren. Als alternatieve instrumenten voor het hardopdenkprotocol planden wij een studentenvragenlijst en een learner-report. De resultaten: Een aantal gesloten vragen in de studentenvragenlijst betroffen leereffecten. Gelet op het gestelde evaluatie-criterium werden deze leereffecten bereikt. In alle gevallen is het oordeel dat men iets heeft geleerd ten aanzien van de probleemaanpak. Voorts is men van oordeel dat na de cursus de eigen aanpak behoorlijk is aangepast in de richting van de SPA. Bovendien is de interesse in de eigen probleemaanpak stevig vergroot. Ook de gerapporteerde leerresultaten in de 'learner reports' verschaffen ondersteuning. In de verdeling van de gerapporteerde leerervaringen komen vooral de demonstreerbare vaardigheden en de algemene inzichten naar voren. Nadere analyse leerde dat het daarbij vooral ging om de Systematische ProbleemAanpak.

ad 3 *Plausibiliteit van andere verklaringen*
Met behulp van de schema's van Campbell en Stanley (1966) en Cook en Campbell (1979) planden we een aantal contrasterende T-toetsen ter controle van alternatieve verklaringen (zie par. 3.2).

Alle toetsingen leverden de verwachte resultaten op, hetgeen betekent, dat de gevonden leereffecten niet konden worden toegeschreven aan dergelijke bronnen.

Voor wat betreft de constructvaliditeit was er wel enigszins sprake van het proefleiderseffect en van instrumentatie. Echter, deze effecten waren zo inherent aan de onderwijssetting dat zij naar onze mening niet tot de foutenbronnen mogen worden gerekend.

5 *Discussie*

Het geheel overziende is onze eindconclusie dat de ontwikkelde cursussen effect hebben op de omvang van de systematiek in de probleemaanpak van de onderzochte studenten. Het effect blijft wel beperkt tot het oriënterende deel van de probleemaanpak. Op de volgorde in de systematiek lijken de cursussen geen effect te hebben.

Een controle op de ondersteuning van bovenstaande claim en de geldigheid van eventuele alternatieve verklaringen leverden geen ontkenning op van de eindconclusie.

Mettes en Pilot (1980) kwamen tot effectieve leerresultaten in cursussen Thermodynamica en Warmteleer waarin eveneens een Systematische ProbleemAanpak centraal stond. Bovendien werd in grote trekken dezelfde FUNDES-methode voor cursusontwikkeling gebruikt. Naar onze mening is met het onderhavige onderzoek aangetoond dat een SPA-benadering óók effectief kan zijn in het sociaalwetenschappelijk domein. Er is bijgevolg sprake van een empirisch onderbouwde generalisatie van het SPA-cursus-concept. Wèl is het zo dat de vormgeving van de SPA een geheel eigen invulling krijgt, gelet op de aard van het sociaalwetenschappelijk domein en de typen problemen die daarin voorkomen. Wij vermoeden evenwel, dat, als er ook in het natuurwetenschappelijk domein sprake is van een (technische) ontwerpprobleem of een verklaringsprobleem (in plaats van de specificatieproblemen van Mettes en Pilot), de SPA-vormgeving qua structuur niet veel zal afwijken van die van het sociaalwetenschappelijk domein (zie b.v. Kramers-Pals & Pilot, 1988). Hoewel wij niet van één onderwijsmiddel zou-

den willen stellen dat deze de leereffecten heeft bewerkstelligd – er is immers sprake van een hele cursus als een complexe onafhankelijke variabele – lijkt het ons wel van belang meer precies de wijze van oriënteren ten aanzien van het leren van een probleemaanpak te onderzoeken. Naar onze mening moet hier zowel vooraf als achteraf worden georiënteerd. Of en in welke verhouding dit moet plaatsvinden, is onduidelijk. Uiteindelijk zijn wij na een aantal formatieve cursussen op deze wijze van oriënteren uitgekomen (zie Terlouw, 1987), echter, dit waren geen strikte experimenteersituaties. Bovendien gaf het nog te weinig zicht op het "waarom" van een dergelijke benadering.

Experimenten in een meer gecontroleerde setting kunnen hier wellicht een nader antwoord geven.

Literatuur

Bishop, Y. M. M., S. E. Fienberg & P. W. Holland, *Discrete Multivariate Analysis:* Theory and Practice. Cambridge (Mass.): MIT Press, 1975.

Breuker, J. A., *Availability of knowledge*. Amsterdam: COWO, 1981.

Campbell, D. T. & J. C. Stanley, *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1966.

Chi, M. T. H., R. Glaser & E. Rees, Expertise in problem solving. In: R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Vol. 1. Hillsdale N. J.: Lawrence Erlbaum, 1982.

Cook, T. D. & D. T. Campbell, *Quasi-experimentation*. Chicago: Rand McNally, 1979.

Cronbach, L. J., *Designing evaluations of educational and social programs*. San Francisco: JosseyBass, 1983.

Denzin, N. K., *The research act*. New York: Mc Graw-Hill, 1978.

Erkamp, A. D., *Ervaringsleren; praktijkinformatie voor vorming en onderwijs*. Amersfoort: Studieboek NCOV, 1980.

Groot, A. D. de, Over leerervaringen en leerdoelen. In: *Handboek onderwijspraktijk*, Aflevering 10. Deventer: Van Loghum Slaterus, 1980.

Hoogerwerf, A., Het leren ontwerpen van overheidsbeleid; een handleiding met toelichting. *Bestuurswetenschappen*, 1984, *1*, 4-24.

Jantos, W., Ueber kooperatives Problemlösen. In: W. W. Dawydow, J. Lompscher & A. K. Markowa (Hrgs.), *Ausbildung der Lerntätigkeit bei Schülern*. Berlin: Volk und Wissen, 1982.

Klerk, L. F. W. de, *Onderwijspsychologie*. Deventer: Van Loghum Slaterus, 1983.

Kossakowski, A., H. Kühn, J. Lompscher & G. Rosenfeld, (Redaktionskollegium) *Psychologische Grundlagen der Persönlichkeitsentwicklung im pädagogischen Prozess*. Berlin: Volk und Wissen, 1982.

Kramers-Pals, H. & A. Pilot, *Leren verklaren 1: Over het Feitelijk Handelingsverloop bij een thermodynamisch verklaringsprobleem*. OC-rapport no. 64. Enschede: CT/OC Universiteit Twente, 1988.

Laurillard, D., *Research methods in student learning*. Niet gepubliceerd paper 3rd EARDHE congress Klagenfurt, 1979.

Larkin, J. H., Teaching problemsolving in physics. In: D. T. Tuma & F. Reif (Eds.), *Problemsolving and education*. Hillsdale N. J.: Lawrence Erlbaum, 1980.

Mettes, C. T. C. W. & A. Pilot, *Over het leren oplossen van natuurwetenschappelijke problemen*. Enschede: Universiteit Twente/Onderwijskundig Centrum, 1980.

Nelissen, J. M. C., De theorie van P. Ja. Gal'perin in discussie. *Pedagogische Studiën*, 1980, *57*, 305-322.

Nelissen, J. M. C., *Kinderen leren wiskunde*. Gorinchem: De Ruiter, 1987.

Nelissen, J. M. C. & A. C. Vuurmans, *Activiteit en de ontwikkeling van het psychische*. Amsterdam: Sua, 1983.

Norušis, M. N., *SPSS-X advanced statistics guide*. New York: Mc Graw-Hill, 1985.

Parreren, C. F. van, Onderwijs als vorm van oriëntering van de leerlingen. In: *Handboek onderwijspraktijk*, aflevering 11, Deventer: Van Loghum Slaterus, 1981.

Parreren, C. F. van, *Bedoeld en bepaald. Een handelingstheoretische visie op de psychologie*. Afscheidscollege Universiteit van Utrecht, 1982.

Parreren, C. F. van, *Ontwikkelend onderwijs*. Leuven/Amersfoort: Acco: 1988.

Parreren, C. F. van & M. C. Schouten-van Parreren, *Onderwijsproceskunde*. Groningen: WoltersNoordhoff, 1981.

Podd'jakov, N. N., Probeerhandelingen in het sensomotorisch leren. *Pedagogische Studiën*, 1979, *56*, 460-470.

Ponomarev: geciteerd bij Parreren, C. F. van, Leren denken getoetst. *Pedagogische Studiën*, 1973, *50*, 319328 en 361-374.

Resnick, L. B., Taskanalysis in instructional design: some cases from mathematics. In: D. Klahr (Ed.), *Cognition and instruction*. New York: Lawrence Erlbaum, 1976.

Steehouder, M. F., C. J. M. Jansen, J. L. C. van der Staak & E. T. Woudstra, *Leren communiceren*. Groningen: Wolters-Noordhoff, 1984.

Terlouw, C., *Syllabus Aanpakken van problemen bij het ontwerpen van overheidsbeleid*. Enschede: Universiteit Twente/Faculteit Bestuurskunde, 1982.
Terlouw, C., *Syllabus Aanpakken van problemen bij het ontwerpen van overheidsbeleid*: Stadsvernieuwing. Enschede: Universiteit Twente/Faculteit Bestuurskunde, 1983a.
Terlouw, C., *Syllabus Aanpakken van problemen bij het ontwerpen van overheidsbeleid*: EG-landbouwbeleid. Enschede: Universiteit Twente/Faculteit Bestuurskunde, 1983b.
Terlouw, C., *De FUNDES-procedure in onderwijsontwikkeling*. Evaluatie van een procedure van onderwijsontwikkeling voor het leren probleemoplossen. Enschede: Universiteit Twente/Onderwijskundig Centrum, 1987.
Terlouw, C., *Een handelingstheoretische invalshoek voor ontwerp en evaluatie van probleemoplosonderwijs*. Paper gepresenteerd op het ORD-symposium "Ontwerp en evaluatie van probleemoplosonderwijs", Leiden, mei 1989.
Terlouw, C., C. T. C. W. Mettes & F. Roemers, *Het leren ontwerpen van overheidsbeleid: een gewenst handelingsverloop*. Enschede: Universiteit Twente/Onderwijskundig Centrum, 1981.
Tichomirov, O. K., Informationsverarbeitungsmodelle und psychologische Modelle für Denkprozesse. In: F. Klix (Hrsg.), *Organismische Informationsverarbeitung Zeichenerkennung - Begriffsbildung - Problemlösen*. Berlin: Akademie Verlag, 1974.
Weeren, J. H. P. van, H. Kramers-Pals, F. F. M. de Mul, M. J. Peeters & H. J. Roossink, *Project Electriciteit en Magnetisme*. OC/TN-rapport nr. 49, Enschede: Universiteit Twente/Onderwijskundig Centrum, 1982.
Winer, B. J., *Statistical principles in experimental design*. London: Mc Graw-Hill, 1970.

Curriculum vitae

C. Tertouw (1947), onderwijskundige, is werkzaam aan het Onderwijskundig Centrum van de Universiteit Twente als coördinator van de sectie probleemoplossen & beoordeling complexe vaardigheden. Hij promoveerde in 1987 op 'De FUNDES-procedure in onderwijsontwikkeling. Evaluatie van een procedure van onderwijsontwikkeling voor het leren probleemopossen'. Tot op heden ontwikkelt hij onderwijs voor, doet hij onderzoek naar en publiceert hij over leren probleemoplossen in het hoger onderwijs.

Adres: Onderwijskundig Centrum. Universiteit Twente, Postbus 217, 7500 AE Enschede.

Manuscript aanvaard 7-9-'89

Summary

Terlouw, C. 'Learning to think in the discipline of Public Administration: learning to solve public policy problems.' *Pedagogische Studiën*, 1990, *67*, 179-192.

Two courses were developed for learning to solve policy design problems. The basis of the course design concerned a mini-theory for teaching-learning processes. A Systematic Problem Approach (SPA) was used learning to solve the policy design problems, generalizing the evidence from science courses. The two courses were evaluated in a field experiment with four groups of students. Significant effects were found in three groups concerning the size of the SPA. The criteria were partly met for the presence and sequence of particular systematic actions. The courses proved to be effective. The SPA viewpoint concerning problem solving could be generalized from physical science to social science.

Discussie

Kanttekeningen bij het congres ter afsluiting van het Rotterdamse project Onderwijs en Sociaal Milieu (1969-1988)

J. DRONKERS
Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam

Het begin van een systematisch overheidsbeleid gericht op het terugdringen van ongelijke onderwijsprestaties ligt rond de overgang van de zestiger naar de zeventiger jaren. Nadat o.a. Van Heek (1968) had aangetoond dat de belangrijkste oorzaak van de ongelijkheid van onderwijskansen tussen kinderen uit verschillende sociale milieus niet lag in 'foutieve' keuzen maar in lage schoolprestaties[1] werden er acties en programma's opgesteld om die lage schoolprestaties te bestrijden. Dit waren meestal lokaal gebonden programma's, veelal opgezet door lokale overheden, opkomende schoolbegeleidingsdiensten en universitaire medewerkers. Elke zichzelf respecterende stad kende zo'n programma, maar die van Amsterdam (Van Calcar, 1980), Utrecht (Stokking & de Vries, 1981) en Rotterdam zijn het bekendst geworden. Het Rotterdamse project werd op 5 oktober 1989 als laatste van de drie genoemde afgesloten met een congres waar het eindrapport (Slavenburg & Peters, 1989) van dit project (in de wandeling het O.S.M.-project genoemd) werd gepresenteerd. Indien men de commotie rond het O.S.M.-project zou afmeten aan de rust op dit afsluitingscongres zou men een foutief beeld krijgen. Vanaf het begin hebben er grote spanningen binnen het projectteam bestaan en tussen de programma's uit Amsterdam, Rotterdam en Utrecht bestond het nodige verschil in opvatting over de waarde van elkaars aanpak en evaluatie. En ondanks alle verschillen tussen deze programma's zijn de uitkomsten niet erg verschillend: de effecten van speciale onderwijsprogramma's in de basisschool gericht op het terugdringen van de lage onderwijsprestaties van kinderen uit de lagere sociale klassen zijn voor zover ze al aanwezig zijn, niet krachtig en slechts van korte duur.

Slavenburg, projectleider en vervolgens directeur van de Rotterdamse School Advies Dienst, was bij de presentatie van de resultaten van het O.S.M.-project op het afsluitingscongres ongenadig. Steeds opnieuw benadrukte hij dat de doelen die in 1968 gesteld waren aan het O.S.M. niet gehaald waren. De schoolvorderingentoetsen lagen nog ver onder het landelijk gemiddelde en de verbeteringen in zittenblijven en uitstroom naar voortgezet en speciaal onderwijs wilde men niet toeschrijven aan de effecten van het O.S.M.-project, omdat deze veranderingen zeer goed het gevolg konden zijn van een landelijke trend. Terecht merkte Meijnen, een van de sprekers, dan ook op dat hij bij lezing van het eindrapport de neiging kreeg het project in bescherming te nemen tegen zijn eigen evaluatoren. Deze niets verbloemende evaluatie van eigen arbeid is dan ook een van de grootste verdiensten van het gepresenteerde eindrapport. Helaas kan men dit niet van alle evaluaties van de andere compensatie- en stimuleringsprogramma's zeggen.

Toch had Meijnen gelijk dat dit project tegen zijn eigen evaluatoren in bescherming genomen moet worden. Allereerst omdat zij vast bleven houden aan de oorspronkelijke doelstelling, nl. dat tenminste het landelijk niveau gehaald moest worden. In feite was deze doelstelling vanaf het begin van het O.S.M.-project irreëel en daardoor onverantwoord. Een dergelijke doelstelling had men niet mogen formuleren op grond van de ook toen beschikbare sociologische analyses naar de oorzaken van de ongelijkheid van onderwijskansen van kinderen uit verschillende sociale klassen. Deze overdreven doelstelling lijkt veeleer voort te vloeien uit een te optimistisch speculeren over de betekenis van het onderwijs zoals dat in het eind van de jaren zestig in de mode was. Het proefschrift van de eerste projectleider van het O.S.M.-project Grandia

(1968) is daarvan mogelijk een van de beste voorbeelden. Mede op grond van dit proefschrift maar ook van de toen heersende tijdgeest zijn de overdreven O.S.M.-doelen geformuleerd en zijn ze, ondanks alle waarschuwingen dat ze onhaalbaar zijn, gehandhaafd. Echter als men met een bepaald middel een irrealistisch doel niet bereikt, zegt dat niets over het middel maar wel wat over het doel. Het is van het project-team nobel dat ze niet aan doelverplaatsing heeft meegewerkt. Maar het is dan onjuist om toch nog conclusies te trekken over de effectiviteit van de uitgevoerde onderwijsprogramma's.

Een tweede reden om het O.S.M.-project tegen zijn eigen evaluatoren in bescherming te nemen is dat zij de programma-effecten die zij wel vonden ten onrechte bagatelliseerden. Met name ten aanzien van de taalprogramma's kunnen ze regelmatig positieve, significante resultaten rapporteren. Uiteraard zijn deze programma-effecten kleiner dan bijvoorbeeld de effecten van voorafgaande schoolprestaties maar dat doet niets af aan hun betekenis. Het is uit de literatuur reeds lang bekend dat de effecten van manipuleerbare variabelen gering is in vergelijking met de effecten van niet-manipuleerbare variabelen. Zo heeft de medische wetenschap geen medicijn gevonden tegen de dood, hoewel op de lange termijn allen sterven. Men concentreert zich daar bijna volledig op veel kortetermijndoelen, zoals het vertragen van een ziekteproces door middel van een medicijn of operatie. Niemand zou bovendien willen suggereren dat de medici met deze kortetermijnmaatregelen moeten stoppen en eerst maar eens moeten vaststellen hoe zij de dood de wereld uit kunnen helpen of het verouderingsproces tot stilstand kunnen brengen. Ook zal niemand een operatie of medicijn weigeren met het argument dat hij of zij op de lange duur toch dood gaat. Het kortetermijneffect van medicijnen of operatie aanvaardt men graag ondanks hun uiteindelijk geringe effect. Bij het O.S.M.-project, maar ook bij de andere onderwijsprogramma's, is ongeveer hetzelfde aan de hand. Ongelijkheid in onderwijskansen is in alle moderne samenlevingen[3] een onvermijdelijk verschijnsel dat nauw samenhangt met het functioneren van die samenlevingen. Deze samenlevingen verschillen dan ook erg weinig in de mate van hun onderwijsongelijkheid. Net als bij de dood moet men dus niet de illusie hebben of wekken dat onderwijsprogramma's die ongelijkheid te niet kunnen doen. Maar dat laatste wil niet zeggen dat men geen aantrekkelijke korte-termijnsuccessen zou kunnen boeken, net zo als met medicijnen: men kan de nadelige gevolgen van onderwijsongelijkheid enige tijd uitstellen. Het effect van maatregelen kan in het licht van de andere effecten nooit zeer groot zijn. Om opnieuw bij een medisch voorbeeld te blijven: het effect van de lichamelijke consistutie op het verloop van een ziekte is veel groter dan dat van een medicijn, zonder dat daaruit de conclusie getrokken mag worden dat het medicijn onbruikbaar is. Op grond van de op het afsluitingscongres gepresenteerde resultaten neig ik er toe met name het O.S.M.-taalprogramma als effectief te beschouwen, gegeven de beperkende randvoorwaarden. En wellicht is het project als geheel ook nog wel effectief als men zich niet blind staart op de leerprestaties maar als men ook oog heeft voor de eigen betekenis van de schoolloopbanen, zoals Meijnen terecht opmerkte. Het is immers heel goed mogelijk dat het O.S.M.-project de onderwijsleerprocessen in de scholen zodanig beïnvloed heeft dat leerkrachten en ouders een andere gedrag t.a.v. door- en uitstroom van leerlingen hebben gekregen. Het verband tussen leerprestaties en schoolloopbaan is immers niet zo absoluut en het verschilt bovendien nog weleens per sociale klasse en school. Het is derhalve jammer dat de O.S.M.-evaluatoren ook op het afsluitend congres zo krampachtig vast hielden aan de oorspronkelijke irreële doelstellingen waardoor ten onrechte het beeld is onstaan dat het O.S.M.-project mislukt zou zijn.

De O.S.M.-programma's zouden bovendien nog wel eens meer effect kunnen hebben dan de nieuwe panacee die door de politieke sprekers op dit afsluitingscongres (Simons, de toenmalige wethouder onderwijs Rotterdam; Wallage, het toenmalige tweede kamerlid; De Wilde, directeur-generaal basisonderwijs) naar voren werd geschoven: 'de effectieve school'. Als men de thans bekende resultaten van het Nederlands effectieve scholen onderzoek (Scheerens, 1989) vergelijkt met die van de O.S.M.-programma's zijn de laatst genoemde het meest effectief. Gelukkig werd dit tijdens de forumdiscussie vanuit de zaal door Tesser opgemerkt maar door het forum niet

opgepakt. Ook werd door de politieke sprekers ten onrechte verondersteld dat de effectieve school juist goed zou zijn voor leerlingen uit de lagere sociale klassen, omdat zij het meest gebaat zouden zijn bij gestructureerd onderwijs. Tevergeefs probeerde Meijnen op het verdelingsaspect van onderwijsongelijkheid te wijzen, waardoor effectieve scholen ook voor leerlingen uit de hogere sociale klassen gunstig kunnen uitpakken en dus de ongelijkheid in het geheel niet worden verminderd. Het O.S.M.-eindrapport maakt zich wellicht onbedoeld mede schuldig aan het bewierroken van de effectieve school, omdat het de gevonden O.S.M.-programma-effecten onvoldoende in de context van de bereikbare effecten hebben geplaatst.

Het is te hopen dat over enige tijd van alle onderwijsprogramma's, die gericht waren op het terugdringen van lage schoolprestaties van kinderen uit de lagere sociale klassen, een meer evenwichtige analyse zal worden gemaakt van de al dan niet bereikte effecten.

Noten

1. Achteraf bleek Van Heek niet gelijk te hebben. Ook een substantieel deel van de ongelijkheid van onderwijskansen zit in 'foutieve' keuzen gedurende de schoolloopbaan. Bovendien vergat Van Heek maar ook zijn critici dat in die tijd de achterstand bij meisjes wel te wijten was aan 'foutieve' keuzen.
2. Nederlandstalige leerlingen uit de lagere sociale milieus moeten als gevolg van de O.S.M.-programma's ten minste de landelijke gemiddelde scores op genormeerde schoolvorderingentoetsen voor rekenen en taal halen en als gevolg daarvan moet het landelijk gemiddelde uitstroomcijfer naar het voortgezet onderwijs, het landelijk gemiddelde zittenblijverspercentage en het landelijk gemiddelde uitstroompercentage naar het buitengewoon onderwijs in de medewerkende scholen gehaald worden.
3. Dat geldt zowel voor zgn. kapitalische als zgn. socialistische landen.

Literatuur

Calcar, C. van, *Innovatieproject Amsterdam. Eindverslag: een opening.* Amsterdam: Van Gennep, 1980.

Grandia, J.H.N. *Uitdaging en antwoord.* Een sociaal-pedagogisch onderzoek naar milieuomstandigheden en maatschappelijke factoren, die een optimale ontwikkeling der talenten hebben verhinderd bij, inmiddels volwassen geworden, jeugdige ongeschoolde arbeidskrachten. Dissertatie Rijksuniversiteit Utrecht: 1968.

Heek, F. van, *Het verborgen talent; milieu, schoolkeuze en schoolgeschiktheid.* Meppel: Boom, 1968.

Scheerens, J., *Wat maakt scholen effectief?* Samenvatting en analyse van onderzoeksresultaten. Den Haag: SVO, 1989.

Slavenburg, J.H. & T.A. Peters, *Het project onderwijs en sociaal milieu: een eindbalans.* Rotterdam: Rotterdamse Schooladvies Dienst, 1989.

Stokking, K.M. & A.K. de Vries, *Een luis in de pels.* Eindverslag van het GEON-project. Den Haag: 1981.

Discussie

Criteria en evaluatie, een moeilijke combinatie

J. H. SLAVENBURG
Rotterdamse School Advies Dienst, Rotterdam

De kern van Dronkers kritiek op de O.S.M.-evaluatie wordt gevormd door de te hoog gestelde doelen, in casu het presteren op ten minste het landelijk gemiddeld niveau op genormeerde schoolvorderingentoetsen. Uiteraard als gevolg van de elf O.S.M.-programma's voor onderwijsstimulering en gezinsactivering. Als de doelen niet zo hoog gesteld waren, zouden de programma's wèl als effectief beschouwd kunnen worden (?), in het bijzonder het taalprogramma. De achterstanden van kinderen uit lagere sociale milieus zullen echter niet volledig weggewerkt kunnen worden, er bestaat geen medicijn tegen de dood.

De critici van de O.S.M.-evaluatiecriteria (de 'te hoge doelen') – Dronkers is niet de enige – realiseren zich kennelijk niet dat, in een bepaald opzicht, deze criteria helemaal niet zo hoog gegrepen zijn. Zij kunnen evenmin goed bruikbare alternatieven aandragen. Ik zal deze twee punten wat nader uitwerken en op basis van deze uitwerking *proberen* een wèl bruikbaar alternatief te formuleren.

De onderwijsachterstand van kinderen uit lagere sociale milieus is per definitie een achterstand ten opzichte van kinderen uit de midden- en hogere milieus. De landelijke gemiddelden op schoolvorderingentoetsen zijn mede gebaseerd op de prestaties van kinderen uit deze lagere milieus; zij zijn daardoor in belangrijke mate lager dan de gemiddelden voor de midden- en hogere milieus. De keuze van landelijke gemiddelden als criteria voor de O.S.M.-evaluatie impliceert dus al (bij realisering door de O.S.M.-programma's) dat de onderwijsachterstanden van leerlingen uit lagere sociale milieus ten opzichte van die uit de andere milieus niet weggewerkt zouden zijn, alleen verminderd. Ik wil maar zeggen, zó hoog gegrepen is het criterium van de landelijke gemiddelden, in dit opzicht, nu óók weer niet. Het realiseren daarvan zou dan ook bepaald nog niet het medicijn tegen de dood geweest zijn.

Dronkers noemt in zijn commentaar geen concrete alternatieven; wèl stelt hij dat sociologische analyses (met betrekking tot de oorzaken van onderwijsongelijkheid) duidelijk gemaakt zouden hebben dat de door O.S.M. gehanteerde evaluatiecriteria niet geformuleerd hadden mogen worden. Naar mijn mening kunnen sociologische noch andere causale analyses over dit onderwerp duidelijkheid verschaffen. Centraal staat de vraag of bepaalde interventies tot effect hebben dat de kansenongelijkheid verminderd of weggewerkt kan worden. Daar is maar op één manier achter te komen: via experimenteren. Sociologische en andere causale analyses kunnen laten zien welke oorzaken beïnvloed moeten worden, zij kunnen echter niet bij voorbaat aangeven hoe groot (of klein) het effect van een interventie zal zijn. Op basis van de effect-evaluaties van deze interventies kan uiteindelijk gesteld worden of de onderwijsongelijkheid verminderd kan worden of niet; niet op basis van causale analyses. Kortom, de sociologische analyses bieden geen aanknopingspunten voor het formuleren van evaluatiecriteria; de tot nu toe ondernomen effect-evaluaties van interventies bieden die wellicht wèl. Ik kom daarop nog terug.

Andere alternatieven die wèl genoemd worden voor het formuleren van evaluatiecriteria (bijv. Meijnen, 1989), die wèl als realistisch gezien worden (en die dus, in de O.S.M.-situatie, lager moeten uitkomen dan de landelijke gemiddelden) worden gebaseerd op delen van standaarddeviaties. De na te streven verbetering is dan bijv. éénderde of éénvijfde deel van een standaarddeviatie (of nòg een ander deel daarvan; in het algemeen bevinden zich de aanbevolen waarden tussen de al genoemde éénderde en éénvijfde). Dit type criteria toegepast in de O.S.M.-situatie zou

voor de lagere leerjaren in het basisonderwijs in het algemeen leiden tot *hogere* na te streven waarden dan de landelijke gemiddelden en in de hogere leerjaren tot lagere. Maar ook deze lagere criteria (in de hogere klassen) zouden door de O.S.M.-programma's niet gerealiseerd worden, terwijl de hogere waarden in de lagere klassen so-wie-so niet gehaald worden. Kortom, deze meer 'realistische criteria' zouden niet leiden tot wèl effectieve O.S.M.-programma's.

Een veel belangrijker dan het zojuist besproken probleem betreft de substantiële inhoud van de standaarddeviatie-criteria: wat stellen zij onderwijsinhoudelijk voor. Voor een antwoord op deze vraag heb ik gekeken naar de O.S.M.-resultaten (uitsluitend voor die toetsen waarvoor de standaarddeviaties van de normeringsgroepen beschikbaar waren). De conclusie is dan nogal eenvoudig: er moet (afhankelijk van de toetslengte, die varieert van 27 tot 50 items, en het feit of men éénvijfde of éénderde standaarddeviatie verbetering nastreeft) een gemiddelde verhoging van de toetsscores plaatsvinden met '1 tot 3 items meer goed' (de in groep 8 BaO gebruikte toets heb ik niet in aanmerking genomen, omdat daarvan geen standaarddeviatie voor de normeringsgroep bekend is; cf. Slavenburg, 1989; fig. 3.7.). Wij zouden dus een kostbare, langdurige interventie moeten opzetten, teneinde een gemiddelde prestatieverbetering van 1 tot 3 punten te bereiken. Dat is een zeer geringe verbetering; iedereen die wel eens een poosje onderwijs gegeven heeft, weet dat – in principe en in de praktijk – aanzienlijk grotere verbeteringen mogelijk kunnen zijn. Dit zijn niet het soort criteria waardoor subsidiegevers en onderwijsgevenden geïnspireerd zullen worden. Dit type effect-criteria is in feite ontleend aan de statistische toetsingstheorie. Daarin zijn effecten onwaarschijnlijke voorvallen; deze voorvallen kan men in een kansuitspraak uitdrukken, maar natuurlijk ook in een (deel van een) standaarddeviatie. Een klein effect is dan éénvijfde van een standaarddeviatie; een groot effect éénderde van een standaarddeviatie. (Gegeven een bepaalde kansverdeling van een gemiddelde kan een en ander ook als een waarschijnlijkheidsuitspraak geformuleerd worden.) Welnu, een klein effect kan van groot praktisch belang zijn en een groot effect kan van weinig praktische betekenis zijn.

Wij hebben hier te maken met de aloude discussie over statistische significantie en substantiële of praktische significantie. Deze discussie is overbekend, ik ga er daarom niet verder op in.

Dit alles leert ons dat vooruitgang in toetsprestaties (met 1, 3 of hoeveel punten dan ook) op zich niet al te veel zegt als daaraan geen inhoudelijke betekenis wordt gegeven. Deze betekenis kan niet ontleend worden aan delen van standaarddeviaties; deze betekenis moet aan andere zaken ontleend worden, zoals aan belangrijke elementen uit de schoolloopbaan van de leerlingen (bijv. doorstroom naar het voortgezet onderwijs en uitstroom naar het speciaal onderwijs). Gezien de sterke relatie tussen schoolprestaties zoals gemeten met genormeerde schoolvorderingentoetsen en de beide genoemde voorbeelden van belangrijke elementen, ligt het voor de hand dat de inhoud van de criteria daaraan zijn betekenis ontleent. Een substantiële verbetering bijv. van de doorstroom naar de hogere vormen van voortgezet onderwijs impliceert een substantiële verbetering van de toetsprestaties. Laten we zeggen dat deze verbetering zo groot mogelijk dient te zijn.

Het volgende probleem is dan: hoe operationaliseren wij 'zo groot mogelijk'? Dat is weer niet te operationaliseren los van onderwijsinterventies. Van deze interventies zou een progressieve verbetering van jaar tot jaar te verwachten moeten zijn; op basis van deze verwachte progressie kan de vooruitgang van de leerlingen gemeten worden (aangenomen dat de te gebruiken toetsen aan de technische eisen daarvoor voldoen en er zich geen belangrijke wijzigingen in de experimentele populatie voordoen). De O.S.M.-programma's zouden bijv., op basis van de evaluatieresultaten, jaarlijks bijgesteld moeten worden; dan kan gemeten worden of progressie in de richting van – laten we zeggen – het landelijk gemiddelde plaatsvindt en zal op zeker moment dit gemiddelde gerealiseerd worden; daarna kan een volgend criterium geformuleerd worden, bijv. het gemiddelde van de middenklasleerlingen. Kortom: een voortdurende effectevaluatie van de interventie, gevolgd door bijstelling van de interventie, weer effectevaluatie enz., net zolang totdat een bepaald

stopcriterium (zoals het landelijk gemiddelde) bereikt is.

Het grote voordeel van deze werkwijze – die dus alleen mogelijk is als daarvoor een jaarlijks bij te stellen interventie beschikbaar is – is dat echte voortgang gemeten kan worden. De interventie wordt pas stopgezet als aan een duidelijk criterium (bijv. de landelijke gemiddelde toetsprestaties) wordt voldaan. Een duidelijk criterium inspireert om met de interventie en de bijstelling ervan door te gaan totdat het daadwerkelijk gerealiseerd is. Het criterium moet zo gekozen worden dat er gunstige gevolgen aan verbonden zijn voor de schoolloopbaan van de leerlingen. Het lijkt me dat iedereen met een dergelijke criteriumformulering zou kunnen leven. Jammer alleen dat bedoelde systematische interventie met jaarlijkse bijstelling nergens in Nederland plaatsvindt. Zolang een dergelijke experimentele en progressieve werkwijze niet wordt geïmplementeerd, wordt de oplossing van het probleem van de kansenongelijkheid niet dichterbij gebracht.

Literatuur

Meijnen, G. W., *Het O.S.M.-project mislukt?* Amsterdam: Stichting Centrum voor Onderwijsonderzoek, 1989.

Slavenburg, J. H., Implementatiewaarde, doeleffectiviteit en transfereffectiviteit stimuleringsen activeringsprogramma's. In: J. H. Slavenburg & T. A. Peters (red). *Het project Onderwijs en Sociaal Milieu.* Rotterdam: Rotterdamse School Advies Dienst, 1989.

Boekbespreking

P. N. Karstanje, *Beleidsevaluatie bij controversiële onderwijsvernieuwing*, Stichting Centrum voor Onderwijsonderzoek, Amsterdam, 1988, 300 pp., *f* 30, –, ISBN 90 6813 196 6.

Het bovengenoemde proefschrift bestaat uit een literatuurstudie over beleidsvorming, innovatie en evaluatie, waaruit een probleemstelling voor een analyse van cases (innovatieprojecten) wordt gedestilleerd. Vervolgens worden de cases beschreven en volgt een beantwoording van de centrale probleemstelling en daarvan afgeleide verbijzonderingen. De algemene probleemstelling luidt: "In hoeverre en op welke wijze houdt de functie van onderzoek en evaluatie in grootschalige onderwijsvernieuwingsprojecten verband met de aard van de beleidsvorming?"

Karstanje kijkt daarbij in zijn literatuurstudie niet alleen naar formele kenmerken van de beleidsvorming (verschillen tussen beleidsactoren per innovatieproject, het meer of minder controversiële karakter van het beleid, en de incrementele dan wel synoptische inslag van het beleid), maar ook naar het onderwerp van de innovatie. Zijns inziens is het van veel betekenis of de veranderingsdoelen "beperkt blijven tot het onderwijs", "dan wel betrekking hebben op buiten het onderwijs te lokaliseren doelen, zoals bijvoorbeeld sociale ongelijkheid in de samenleving of de relatie onderwijs-arbeidsmarkt".

Bij de nadere typering van onderzoek en evaluatie in het kader van innovatieprojecten belicht Karstanje onder meer de functie van evaluatie (een legitimerend-controlerende functie versus een communicatieve 'formatieve' functie) en het probleem van het niet eenduidig zijn van evaluatie-criteria en -standaarden, in een situatie waarin er meerdere belanghebbenden zijn bij een evaluatie. Verder gaat hij in op de evaluatiemethodologie, waarbij hij de (quasi)experimentele of 'klassieke' benadering plaatst tegenover meer holistische en naturalistische benaderingen.

Ten slotte wordt in het literatuurgedeelte aandacht besteed aan de politieke rol die evaluatie-onderzoek zou kunnen spelen. Zoals ook in eerdere publikaties van zijn hand (Karstanje, 1984) toont de auteur zich ook in zijn dissertatie bezorgd over de status quo bevestigende invloed van 'klassiek' evaluatie-onderzoek.

Uit de belangrijkste dimensies van innovatie, beleid en evaluatie die in het literatuurgedeelte behandeld worden, worden aandachtspunten afgeleid voor de informatieverzameling bij een drietal cases. De drie cases zijn het project middenschool en voortgezet basisonderwijs, de vernieuwingen in het middelbaar beroepsonderwijs en het mavo-project. Elk van de cases wordt beschreven aan de hand van een aantal fasen in de wordingsgeschiedenis en het verloop van de innovatieprojecten, waarbij systematisch de inbreng van verschillende betrokken instanties ('actoren') wordt geïnventariseerd. De toegepaste methode van gegevensverzameling is inhoudsanalyse van allerlei relevante documenten (notulen van vergaderingen, onderzoeksrapporten e.d.). Na de beschrijving van de drie projecten en het reilen en zeilen van evaluatie daarbinnen worden de drie projecten onderling vergeleken.

De voornaamste conclusies die deze vergelijking oplevert, zijn (beknopt samengevat) de volgende:

- "het meest bepalende onderscheid tussen de drie projecten qua innovatie-onderwerp is de mate waarin de innovatie het karakter van structuurverandering kreeg" (p. 261);
- "naarmate het domein van de veranderingsdoelen minder binnen het onderwijs ligt en meer een structuurverandering betreft, wordt de complexiteit groter, neemt de controversialiteit toe en blijkt de beleidsvorming minder synoptisch en incrementeel te zijn" (p. 262);
- "naarmate beleidsactoren meer tegen een structuurwijziging zijn, zijn ze tegen experimenten of kennen ze er een evaluatiefunctie aan toe (ter onderscheid van een ontwikkelfunctie), waarbij evaluatie een summatief vergelijkende functie krijgt en leerprestaties en experimenten moeten worden vergeleken met leerprestaties in het bestaande stelsel" (p. 263);
- "het gebruik van onderzoeksgegevens kan in het onderzochte materiaal nauwelijks worden vermeld" (p. 264).

Op basis van deze uitkomsten komt de auteur tot de aanbeveling om beleidsevaluaties vooraf te laten gaan door contextanalyses en analyses van de visies en belangen van de voornaamste beleidsactoren, bepleit hij een verbetering van de communicatie tussen beleidsactoren en waarschuwt hij tegen evaluatie die als "voorwaarde voor politieke besluitvorming wordt gesteld" (p. 272) in een situatie waarin de besluitvorming "van het sociale interactie-type is".

Het literatuuroverzicht is redelijk up to date wat betreft typeringen van beleidsvoering en evaluatiebenaderingen. Wel worden sommige kwesties wel erg oppervlakkig behandeld, zoals de kwestie van afwegingen betreffende de interne en externe validiteit van evaluatie-onderzoeksdesigns (p. 49). Ook is het

wat teleurstellend dat zowel beleidsvoerings- als evaluatieconcepten uiteindelijk toch weer tot twee paren van extremen worden teruggebracht: resp. rationeel-synoptisch vs. incrementeel en experimenteel vs. holistisch, proces-gericht en formatief evaluatieonderzoek. Gezien de stand van de conceptuele ontwikkeling zijn meer verfijnde en genuanceerde typeringen te gebruiken. Verder is mij opgevallen dat er weinig is gedaan met theorievorming betreffende onderwijsinnovatie.

De vergelijkende case-studiebenadering is voor de probleemstelling in kwestie de meest geëigende – waarschijnlijk ook wel de enig mogelijke – aanpak. Vaak wordt bij dit soort onderzoek aanbevolen om meerdere methoden van data-collectie toe te passen. Wat dat betreft zou de hier aan de orde zijnde studie versterkt kunnen zijn wanneer naast de documentenanalyse bijvoorbeeld ook gesprekken met sleutelfiguren waren uitgevoerd. Uiteraard heeft een vergelijkende case-studiebenadering met een n van 3 cases zijn beperkingen. Kennelijk heeft de auteur het niet nodig gevonden hierover uit te wijden.

Als er in deze studie iets overtuigends wordt aangetoond, dan is het wel de stelling dat de beleidsmatige context van – vooral de meer ingrijpende – onderwijsinnovaties bepaald wordt door meerdere beleidsactoren. Dit schept problemen voor een technisch verantwoorde en bruikbare evaluatie. Karstanjes analyse van het pluriforme krachtenveld rondom evaluaties van innovatieprojecten verschilt onder meer van eerdere analyses, waarin dit ook is gebeurd (Scheerens, 1983; Hofstee, 1984), door een relatieve verwaarlozing van de evaluatie-onderzoekers, c.q. hun institutionele basis, c.q. hun plaats binnen een wetenschappelijk forum, als partij. Als partij die zich kan beroepen op bepaalde professionele principes over hoe evaluaties wel en niet moeten worden uitgevoerd, wel te verstaan, en niet – zoals Karstanje geneigd is de beleidsactoren neer te zetten – als politieke voor- of tegenstanders van de innovatie. Hoewel de onderzoekersgemeenschap, in de projecten die worden beschreven, een betrekkelijk zwakke rol speelde, gaat het te ver om evaluatie-onderzoek *uitsluitend* als speelbal van andere belanghebbende instanties te zien.

Een overtuiging die de auteur al langer uitdraagt – en die hij ook in deze studie bevestigd ziet – is de opvatting dat tegenstanders van innovaties voorstanders van vergelijkend effect-gericht evaluatie-onderzoek zijn en dat eigenlijk ook geldt dat voorstanders van effect-gerichte evaluatie tegenstanders van innovatie zijn. Op p. 54 denkt hij steun te vinden voor zijn opvatting dat evaluatieonderzoek een "establishment oriëntatie" (eerder, 1984, sprak hij van een "conservatieve bias") heeft door een wat uit zijn verband gehaald citaat uit Weiss, 1975, p. 21. Weiss bedoelt met establishment echter de instantie die verantwoordelijk is voor het innovatieprogramma en die dus zeker niet zit te wachten op 'slecht nieuws' over het programma. Bovendien stelt ze enkele regels verderop juist dat evaluatoren meer 'liberal' zijn dan hun opdrachtgevers en daarom in hun onderzoek eerder geneigd zijn tot wat men in dit verpolitiekte jargon zou moeten aanduiden als 'progressieve bias'.

Aan deze opvatting van Karstanje moest ik denken toen ik onlangs een directeur van een middenschool op de televisie hoorde klagen over het feit dat de overheid er kennelijk niet in geïnteresseerd was geweest om al het goeds dat op zijn school bereikt was met behulp van onderzoek vast te stellen.

Voor een kritiek op de stelling dat aan effectgericht evaluatie-onderzoek een politiek conservatieve vertekening inherent is, zij verwezen naar Scheerens, 1984a, 1984b. Het is ook gemakkelijk om tegenvoorbeelden uit de Nederlandse evaluatiepraktijk te noemen: hoe zit het met de conservatieve bias in de nota bene *interne* effect-evaluatie van het OSM-project? En hoe is het te verklaren dat in thans op stapel staande en deels in uitvoering zijnde evaluatie-projecten als de evaluatie van het onderwijsvoorrangsbeleid en de evaluatie van de basisvorming vergelijkende effectmetingen een centrale plaats innemen? Toch zeker niet omdat de belangrijkste beleidsactoren zo tegen deze onderwijsvernieuwingen zijn!

Het tekort doen aan de wetenschappelijke integriteit van evaluatie-onderzoek dat spreekt uit de bovengenoemde opvattingen, brengt met zich mee dat de aanbevelingen om de praktijk van beleidsgericht evaluatie-onderzoek te verbeteren nogal mager zijn. Naast de zinvolle aanbeveling om evaluatie-onderzoek door context-analyses vooraf te laten gaan (maar dat is ook al eerder voorgesteld) worden eigenlijk alleen aanbevelingen van Van de Vall herhaald, daarbij voorbij gaand aan de discussies waartoe diens combinatie van 'epistemologische' en 'implementaire' validiteit aanleiding heeft gegeven.

Kortom, het proefschrift in kwestie is een interessante studie van de beleidscontext van onderwijsinnovatieprojecten in de periode van ca. 1975 tot ca. 1985, maar kennelijk leidt de optiek van een onderwijsinnovator, die zich klassieke en 'conservatieve' evaluatoren van het lijf wil houden, tot weinig nieuwe gezichtspunten voor de uitvoeringspraktijk van beleidsgericht evaluatie-onderzoek.

J. Scheerens

Literatuur

Hofstee, W. K. B., Evaluatie: beoordelen van kwaliteit. *Tijdschrift voor Onderwijsresearch*, 1985, *10*, 285-288.

Karstanje, P. N., Evaluatie bij grootschalige onderwijsinnovaties. *Pedagogische Studiën*, 1984, *61*, 127-138.

Scheerens, J., De zogenaamde conservatieve bias bij enkele gangbare evaluatie-designs. *Tijdschrift voor Onderwijsresearch*, 1984a, *9*, 161-170.
Scheerens, J., Evaluatie-onderzoek en politieke bias. *Kennis en Methode*, 1984b, *8*, 41-52.
Scheerens, J., *Evaluatie-onderzoek en beleid. Methodologische en organisatorische aspecten*. Harlingen: Flevodruk B.V., 1983.
Weiss, C.H., Evaluation research in the political context. In: E.L. Struening & M. Guttentag, *Handbook of Evaluation Research* (13-26). Beverley Hills: Sage, 1975.

B. Lumer-Henneböle & E. Nijssen, *Basisschulen in den Niederlanden*. Böhlau, Köln/Wien, 1988, 215 pag., ISBN 3-412-065889.

Wat denkt het buitenland over ons eigen onderwijssysteem? Het antwoord op deze vraag blijft voor velen fascinerend. In het voorliggend boek kijken Duitse ogen naar de nieuwe Nederlandse basisschool als innovatieproces. Systematische vergelijkingen met de BRD worden niet beoogd. De titel dekt de lading. De lezer krijgt een degelijk verslag aangeboden dat gebaseerd is op schoolbezoeken, lectuur en gesprekken met diverse Nederlandse onderwijsbetrokkenen.

Het verwondert de auteurs dat het Nederlandse onderwijs en zijn innovaties weinig bekend zijn in de BRD. Wat de basisschool betreft werden in Nordrhein-Westfalen analoge hervormingen doorgevoerd zonder enige verwijzing naar Nederland. Onder meer deze verwondering motiveerde de auteurs eerst tot praktijkverkenning en daarna tot meer diepgaande theoretische studie en analyse van het Nederlandse onderwijssysteem, meer specifiek bevraagd op de conceptualisering en implementatie van de achtjarige basisschool. Na de rapportering van een schoolbezoek (hoofdstuk 1) wordt de aandacht gericht op Nederlandse uitgangspunten die voor buitenlanders ongetwijfeld moeten opvallen. Verzuiling en onderwijsvrijheid worden geduid als basisprincipes van het Nederlandse onderwijssysteem. Deze uitgangspunten worden beknopt gesitueerd op een historische en socio-economische achtergrond. De impact op onderwijstheoretische grondslagen wordt aangetoond (hoofdstuk 2). Een derde hoofdstuk behandelt de globale onderwijsstructuur en concentreert zich verder op de sociale selectiefunctie van het Nederlandse onderwijs. Niet slechts is de kijk van auteurs Duits, maar ook vrouwelijk. Vermoedelijk is daarom ook het vierde hoofdstuk gewijd aan de kansen(on)gelijkheid van (jongens en) meisjes in het Nederlandse onderwijs.

In de daarop volgende hoofdstukken (5, 6, 7) wordt het hoofdthema van het boek centraal gesteld, nl. het concept, de inhoud en de institutionalisering van de nieuwe Nederlandse basisschool. Schoolwerkplan, zorgverbreding, pedagogisch klimaat, lerarenopleidingen, schoolbegeleidingsdiensten, Stichting voor Leerplan-Ontwikkeling, en andere typisch Nederlandse instituten, worden vertalend geïnterpreteerd voor en door Duitse geïnteresseerden. In een slothoofdstuk becommentarieert en onderbouwt Prof. K. Doornbos deze Duitse kijk op het Nederlandse onderwijs.

Het concept en de inhoud van dit boek (weliswaar bescheiden in opzet en drukkwaliteit) verdienen aanbeveling. Buitenlandse auteurs voeren de lezer, via een verkenning van de dagelijkse onderwijspraktijk, binnen in een aantal algemene maar (voor buitenlanders) zeer belangrijke kenmerken van het Nederlandse onderwijsbestel, om uiteindelijk inzicht te verlenen in de basisschool als innovatieproces. Daarbij wordt gerefereerd naar en gesteund op zowel onderwijswetgeving als op - statistieken en theorieën die relevant zijn voor het Nederlandse onderwijs.

Uiteraard is dit boek interessant voor Duitse beleidsverantwoordelijken inzake onderwijs. Maar ook voor Nederlanders kan het verrijkend zijn hun onderwijssysteem te bekijken door een Duitse bril.

W. Wielemans

Swaan, Abram de, *Zorg en de staat. Welzijn, onderwijs en gezondheidszorg in Europa en de Verenigde Staten in de nieuwe tijd*. Vertaald door H. Hendriks en herzien door A. de Swaan. Uitgeverij Bert Bakker, Amsterdam, 1989, 340 pag., paperback *f*45,-; gebonden *f*64.50; ISBN 90 351 0685 7 en 90 351 0701 2.

De Swaan heeft een magnifiek boek geschreven, dat wel eens een bestseller zou kunnen worden. Het is ouderwets in de goede zin van het woord. De auteur ontpopt zich in een aantal opzichten als een twintigste eeuwse 'philosophe uit de achttiende eeuw', hetgeen in een tijd van post-modernisme, superspecialisme en dominantie van het neo-klassieke economisch denken van durf getuigt en verademend werkt. Het onderwerp is breed, het betreft de ontwikkeling van een scala van instituties in verschillende sectoren van zorg in vijf landen gedurende vijf eeuwen. Het werk is opgezet vanuit een bepaalde visie op het civilisatieproces en uitstekend gecompo-

neerd. Bovendien kan De Swaan schrijven. Bij het lezen moet men af en toe zelfs glimlachen. Het leven in de negentiende eeuwse grote stad wordt in hoofdstuk vier wel zeer beeldend beschreven en het slot van het laatste hoofdstuk getuigt van bijkans eschatologisch aandoende visionaire kracht. De erudiete kennis en de (soms wat populaire?) wetenschappelijke inzichten uit meerdere disciplines zijn schijnbaar moeiteloos in de tekst en de verschillende sectorale cases verwerkt. Ook de talrijke neologistische vondsten zijn functioneel, bijvoorbeeld het rekbare begrip 'metropolitane' elites. Op pag. 147 holt de auteur zichzelf evenwel voorbij waar hij met een sweeping statement concludeert, dat 'het watercloset de rechtstreekse verbinding werd tussen de individuele burger en de collectiviteit'. Verder is alleen de eerste zin van de tweede alinea op pag. 241 onhelder.

Vanuit onze achtergrond bekeken we het boek primair vanuit het perspectief van een (sociaal) pedagoog die zich in het kader van de historische en vergelijkende onderwijskunde vooral bezig houdt met beroepsonderwijs-scholing-en-training en die belangstelling heeft voor (onderwijs-)economie. Vandaar dat onze aandacht vooral uitging naar het sectorale thema onderwijs in hoofdstuk drie dat handelt over het ontstaan van nationale communicatienetwerken en de verbreiding van het verplicht lager onderwijs op nationaal niveau in de loop van de achttiende en negentiende eeuw. De andere onderwerpen betreffen het ontstaan en de ontwikkeling van respectievelijk de armenzorg (van liefdadigheid naar armenhuizen), de gecompliceerde stelsels van sociale zekerheid, de stedelijke gezondheidszorg (gelieerd aan de aanleg van gemeentelijke waterleidingen en rioleringen) en de opkomst en ondergang van de onderlinge arbeidersverzekering welke een centraal thema van het boek, de dilemma's van collectieve actie (met name het basismechanisme van het 'dilemma van de gevangene' en de 'zwartrijder') wel zeer schrijnend illustreert. Het eerste hoofdstuk bevat de wetenschappelijke verantwoording. Hier levert de auteur onder andere kritiek op de welzijnseconomie en de geschiedkundige sociologie en komt vervolgens tot zijn eigen synthetische sociogenetische benadering (pag. 14). Het slothoofdstuk, de apotheose van het boek, combineert (hier weergegeven in de bewoordingen van de auteur, pag. 22) de thema's in een samenvatting en een bespreking van de ontwikkeling in de verzorgingsstaten van 1945 tot heden. Daarna komt de transformatie van de middenklasse aan de orde, en de gevolgen van bureaucratisering en professionalisering voor de cognitieve oriëntatiemiddelen in het dagelijks leven (proto-professionalisering), de affectbeheersing (het civilisatieproces in engere zin) en de morele instelling (de wording van het sociaal bewustzijn). De Swaan is sterk beïnvloed door het denken van Norbert Elias, hetgeen hij ook nadrukkelijk vermeldt, maar veel te eigenzinnig en creatief om tot epigonendom te vervallen.

De historische en vergelijkende benadering creëert afstand en in samenhang met de door de auteur aangehangen visie eveneens zicht op institutionele ontwikkelingen binnen de sectoren (inclusief het onderlinge verband daartussen), vooral doordat de auteur wijst op ten grondslag liggende meer algemene mechanismen. Ook in de gedetailleerde beschrijvingen attendeert De Swaan ons op overeenkomsten binnen en tussen de als exemplarische casestudies te beschouwen sectorale thema's en de daarbij behandelde landen. Voor wie zich direct of indirect op enigerlei wijze bezig houdt met het onderwerp onderwijs-arbeid wordt het verband tussen het ontstaan en de ontwikkeling van het onderwijs en de ontwikkeling van instituties binnen de andere sectoren verhelderd. Door het scherp observeren, de humanistische openheid van geest en de reeds gememoreerde taal-virtuositeit verworden de macro- en micro-sociologische verklaringen niet tot schematismen. Integendeel, de auteur houdt steeds ook oog voor nuances en verschillen (van dit laatste getuigen onder andere de landenstudies en het uitgebreide notenapparaat, pag. 263-306). Wel past alles in een geheel. Die 'holistic view' is meeslepend, prikkelt op onderdelen bij enig nadenken tot tegenspraak en werpt af en toe op bepaalde (actuele en oude) vraagstukken enig nieuw licht, waardoor de lezer in een enkel geval min of meer gedwongen wordt de eigen zienswijze bij te stellen. Dit gold voor ons bij voorbeeld voor de verhouding tussen taal als code voor nationale communicatie en de ontwikkeling van het onderwijs, die tussen kerk en staat (inclusief de verzuiling in ons land) en het verband tussen bepaalde vormen van sociale zekerheid en delen van het onderwijs. Verhelderend is voorts het gebruik van begrippen en inzichten uit de economie, bij voorbeeld: surplus, collectief goed, overdrachtsvermogen, externe effecten (in verband met de probleemstelling alleen de negatieve?) en het reeds vermelde populaire freerider- en prisoners-dilemma. Deze laatste twee ontleningen uit de 'sombere wetenschap' der economie (en politicologie) getuigen in hun toepassing in dit boek van twintigste eeuws realiteitsbesef, scepticisme en een zekere afstandelijke gelatenheid bij de (in wezen achttiende eeuwse?) humanist De Swaan. Steeds weer treden in het civilisatieproces onbedoeld negatieve externe effecten op die kunnen worden voorkomen en bestreden doordat mensen en groepen elkaar vertrouwen of overreden. Uiteindelijk blijkt echter steeds weer geestelijk en vooral centraal gezag vereist. De maatschappelijke ontwikkeling verloopt mede daardoor in de richting van een steeds hoger en subtieler niveau van integratie, op macro-niveau van het lokale en regionale naar het nationaal niveau en binnenkort waarschijnlijk het mondiale. De statenvorming en collectivisering van zorg verklaart De Swaan met behulp van een vierzijdige figuratie tussen vier strijdende partijen die wisselende coalities

vormen, waarbij het regime, de metropolitane elites, de eerste viool spelen en de kleine burgerij uiteindelijk het onderspit delft. Wordt de invloed die grote figuren als Roosevelt, De Gaulle (1958) en vooral Bismarck ook volgens de auteur hebben gehad hier indirect niet wat onderschat of zijn deze staatsmannen toch louter produkt van plaats en tijd?

Op de omslag van het boek is het schilderij De Barmhartige Samaritaan van Van Gogh afgebeeld. De Swaan heeft oog voor de harde werkelijkheid, doch beseft tegelijkertijd zeer wel dat bij het voortschrijden van het civilisatieproces ook onvermijdelijk zaken verloren gaan. Dit gevoel komt het scherpst tot uiting bij de behandeling van de teloorgang van streektalen en randtalen als het Fries en Welsh. (Biedt de internationalisering in Europa en de tendens tot culturele pluriformiteit juist thans deze kleine cultuurgemeenschappen niet weer enige kansen of is dat te voluntaristisch gedacht? M.S.)

Impliciet kiest De Swaan duidelijk voor de verzorgingsstaat. De grote waarde en vooral het diepgeworteld zijn ervan wordt indringend geschetst (pag. 234 t/m 236 en 242). Thatcherisme en Reaganomics blijken grotendeels retoriek te zijn. Opvallend is, dat Illich slechts eenmaal genoemd wordt, pag. 243, en niet in de literatuurlijst vermeld staat. Achterhuis ontbreekt geheel. De auteur gaat uitvoerig in op het probleem van de verzorgingsstaat. Met name op de professionalisering, waarbij ook aandacht aan de sector onderwijs wordt besteed. In tegenstelling tot de hoofdlijn uit het boek, de ontwikkeling van lokaal naar mondiaal, wordt hier voor het oplossen van de uitvoeringsproblemen (van-en-binnen de verzorgingsstaat) een beroep gedaan op het lokale en regionale niveau.

Een fundamentele vraag met mogelijk sociaalpedagogische consequenties is, in hoeverre de mensen de toenemende interdependentie en abstractheid van het maatschappelijk gebeuren aankunnen, ja in hoeverre ze mee mogen doen (in de Zweidrittel-Gesellschaft). De Swaan is geen Mannheim aan wie recentelijk een thema-nummer van Beleid en Maatschappij gewijd werd. Aanzetten tot actief vormgevend interveniëren, construeren en reconstrueren blijven achterwege, hoewel de inhoud van het boek daartoe in principe wel uitnodigt en aangrijpingspunten biedt. Interessant is of iemand als De Swaan dit wel een taak van de wetenschap(-per) acht.

Het begrip 'welzijn' in de ondertitel werkt verwarrend. Het begrip 'welfare' uit de Engelse versie is te letterlijk vertaald. Het Nederlandse begrip welzijn wordt doorgaans eng, namelijk CRM/-/WVC-achtig, opgevat. In de ondertitel hadden wellicht beter de diverse vormen van zorg of zekerheid kunnen worden opgesomd, zo had ook de sociaal-economische dimensie van de inhoud van het boek in de titel tot uitdrukking kunnen komen. Ook de oorspronkelijke titel 'In Care of the State' vinden we beter dan de Nederlandse 'Zorg en de Staat'. Afgaande op de betiteling van de beide uitgaven blijkt Nederland in overeenstemming met de daarin vervatte theorie al af te zinken naar provinciaal niveau en begeeft De Swaan zich op het mondiale pad.

Sommigen zullen de auteur finalisering verwijten of stellen dat het schrijven van een boek als dit geen wetenschap is. In de wetenschappelijke verantwoording in het eerste hoofdstuk geeft De Swaan wat ons betreft afdoende antwoord op dit soort opmerkingen (met name pag. 9, 12 en 18). En bovendien, een beetje 'verbeelding' mag toch.

Een aantal (onderwijs-)specialisten zal ongetwijfeld opmerken, dat bepaalde zaken vergeten zijn, te globaal behandeld worden of aan de hand van te oude literatuur. Zo ontbreken bij voorbeeld Willis en Bourdieu (Bernstein is wel prominent aanwezig); Archer komt daarentegen slechts terloops aan de hand van een publikatie uit 1971 ter sprake. In hoofdstuk één geeft De Swaan ook op deze kritiek antwoord. Wat de thematiek betreft is de auteur in elk geval wel bij de tijd.

Met betrekking tot niet, of minder regulier onderwijs als het leerlingwezen en bedrijfsopleidingen kan de vraag worden opgeworpen of de collectivisering van dit onderwijs zich enkel trager voltrekt of dat het er anders verloopt, bij voorbeeld ambivalenter.

De Swaan heeft een inspirerend boek geschreven, dat ongetwijfeld een breder publiek dan sociologen en andere wetenschappers zal bereiken. Voor een deel is het ook te beschouwen als een synthese van zijn voorafgaande werk (raadpleeg de literatuur-opgave). In een aantal opzichten is het ook een beetje een Amsterdams boek en De Swaan de professorale lefgozer.

M. Santema

Mededelingen

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
15e jaargang, nr. 1, 1990

In memoriam M.J. Langeveld, door J. Rispens
De veelzijdige pedagogiek van Janusz Korczak, door J.W.A. Berding
De zin van een interpretatieve benadering van het klasgedrag van leerkrachten en leerlingen bij het begin van een nieuw schooljaar, door M. Clement en J.C. Verhoeven
Ouders en leeftijdgenoten in het persoonlijke netwerk van jongeren, door W. Meeus

Tijdschrift voor Onderwijsresearch
14e jaargang, nr. 6, 1989

Schoolkenmerken en leerlingfunctioneren: een replicatie-onderzoek, door A.W.M. Knuver
Een onderzoek naar de effecten van auditieve feedback op leesfouten, door G.W.G. Spaal, P. Reitsma en H.H. Ellerman
Notities en Commentaren
Test-toetsgebruik in programma-evaluatie: herziene richtlijnen, door H.P.G.M. Caminada
Peil van PPON-onderzoek kan hoger, door Y.J. Pijl en E.G. Harskamp
Normaal Gronings peil? Een reactie op Pijl & Harskamp, door Th.J.H.M. Eggen, N.D. Verhelst en J.M. Wijnstra

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 1, 1990

Onderzoek naar effecten van speciaal onderwijs: lastig maar niet onmogelijk, door J.C. van der Wolf en C.M. van Rijswijk
Initiatief en kinderen met motorische en/of verstandelijke beperkingen, door G.P. Loots en F.A. Goossens
Traditie en vernieuwing, door J. Rispens
De loopbaan van lom-leerlingen, door J.C. van der Wolf

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 2, 1990

Ontwikkelingsconcepties van ouders met normale en geestelijk gehandicapte kinderen: overzicht en implicaties, door S.A. de Roos en A.E. Jackson
De effecten van spelbegeleiding in het regulier en het speciaal onderwijs, door S.E.K. Faber
Casus Stephan, door H. Pijnenburg

Ontvangen boeken

Grift, W. van de & C. Kurek-Vriesema, *Schoolleidingen in de basisschool*, Swets & Zeitlinger, Lisse, 1990, *f* 24,—.

Horst, A. van den, I. Brouwer, J. Rensen & K. Twilt, *Slecht horen.* Boom, Meppel, 1990, *f* 27,50.

Huizing, D. *Een schepje er boven op! Over natuur- en milieu-educatie en pedagogiek.* SDU-Uitgeverij, Den Haag, 1989.

Hulst, J.W. van (red.), *Pedagogiek MO in beweging; 65 jaar geschiedenis.* Vereniging tot Bevordering van de Studie der Pedagogiek, Amsterdam, 1990.

Hulst, J.W. van, *Kerkelijk antisemitisme. Enkele voorbeelden uit Nederland en Duitsland 1920-1945.* Uitgeverij Merweboek, Sliedrecht, 1990.

Jong, O. de, *Rekenen aan reaktie. Een didactische studie van chemische berekeningen in begin bovenbouw havo/vwo* (dissertatie). Rijksuniversiteit Utrecht, 1990.

Leemans, E. & M. Bruinsma, *Omgaan met kinderen van 1½ tot 4. Theorie en praktijk.* Wolters Noordhoff, Groningen, 1990, *f* 37,50.

Matthijs, W., *De zelf/ander-representaties van gedragsgestoorde kinderen. Een onderzoek naar interne objectrelaties* (disseratie). Swets & Zeitlinger, Lisse, 1990, *f* 34,—.

Ruijssenaars, A.J.J.M. & J.H.M. Hamers (red.), *Dyslexie. Ernstige lees- en spellingproblemen. Een overzicht van theorie en praktijk.* Acco, Amersfoort/Leuven, 1989, *f* 38,75.

Singley, M.K. & J.E. Anderson, *The transfer of cognitive skill.* Harvard University Press, Cambridge/London, 1989, $ 35.95.

Versloot, A.M., *Ouders en vrijheid van onderwijs. Schoolkeuze in de provincie Utrecht* (dissertatie). Academisch Boeken Centrum, De Lier, 1990.

Het schrijfmodel van Hayes en Flower: een kritische analyse

L. VANMAELE en J. LOWYCK
Katholieke Universiteit Leuven

Samenvatting

Het schrijfmodel van Hayes en Flower wordt frequent benut in onderzoek over schrijven en meer recentelijk ook voor het ontwerpen van computeromgevingen. In deze bijdrage worden de kenmerken van het model kort geschetst om het op zijn waarde te kunnen beoordelen. Eerst komen de doelen, methoden en uitkomsten van het door Hayes en Flower verricht empirisch onderzoek aan de orde. Vervolgens worden de beperkingen van het model geïndiceerd, door een beroep te doen op een aantal aanvullende modellen. Voorts wordt het model beoordeeld aan de hand van de volgende vier criteria: werkelijkheidswaarde, veralgemeenbaarheid, prescriptieve waarde en helderheid. Het artikel wordt afgesloten met een aantal conclusies aangaande de waarde van het model voor het construeren van een leeromgeving voor het schrijven.

Inleiding

In het kader van het onderzoeksproject 'Functionele Informatica in het Beroepsonderwijs en het ontwikkelen van basisvaardigheden in lezen en schrijven" werd door het Onderzoekscentrum voor Educatieve Technologie (Educatec) van de K.U. Leuven gezocht naar een model voor het ontwerpen van een schrijfomgeving voor taalzwakke leerlingen (Lowyck, De Gelder & Vanmaele, 1987). In dit verband is heel wat aandacht uitgegaan naar het model van Hayes en Flower (1980a, 1980b, 1986). Immers, dit probleemoplossingsmodel vormt in recent onderzoek naar schrijfprocessen frequent het uitgangspunt voor o.m. de toetsing ervan aan het eigen onderzoek (Bochardt, 1984), de verdere verfijning van het model (Molitor, 1987), de constructie van een alternatief model (De Beaugrande, 1984; Bereiter, 1980; Scardamalia & Paris, 1985; Scardamalia & Bereiter, 1986, 1987) of het ontwerpen van een computerleeromgeving (Rubin & Bruce, 1984; Van der Geest, 1984). Zelfs onderzoekers die hun aandacht op andere, meer emotionele aspecten richten, maken voor de omschrijving van de cognitieve dimensie eveneens gebruik van het model van Hayes en Flower (zie o.a. Rose, 1980). Het ligt in de bedoeling van deze bijdrage om het model van Hayes en Flower nader te beschrijven en te beoordelen.

Vooraleer tot een kritische analyse van het model zelf te komen, wordt eerst een overzicht gepresenteerd van de belangrijkste kenmerken van het model, evenals van zijn empirische onderbouw en van de achterliggende theoretische kaders (paragraaf 1). Zo'n beschrijving verheldert ongetwijfeld de essentiële componenten van het besproken model en draagt bij tot een ruimere situering.

Vervolgens komt de vraag aan de orde of het door Hayes en Flower ontwikkelde model voor probleemoplossen wel geschikt is voor het vatten van het schrijven als complexe activiteit. M.a.w.: is schrijven wel een kwestie van probleemoplossen. Om hierop een antwoord te formuleren, wordt eerst nagegaan wat onder 'probleemoplossen' kan worden begrepen. Uit deze analyse blijkt, dat voor een goed begrip ervan, de onderscheiden types van problemen mede onderkend dienen te worden. Nadat deze kwestie is verduidelijkt, wordt gezocht of en in welke mate het schrijven als een vorm van probleemoplossen kan worden beschouwd en tot welk probleemtype de schrijfopdrachten behoren. Bovendien wordt naar een verruiming van het model gezocht, o.m. door na te gaan of het schrijven ook vanuit andere invalshoeken kan worden verhelderd, zoals: een model voor de analyse van artistieke en sociale processen, het 'problem-finding'-model en de schematheorieën (paragraaf 2).

De informatie uit de beide vorige paragrafen stelt ons aansluitend in staat om het door Hayes en Flower uitgewerkte model in paragraaf 3 kritisch te beoordelen. Mede gezien de uitgangspunten, met name het ontwerpen van

een adequate schrijfomgeving, wordt het schrijfmodel hoofdzakelijk beoordeeld op de volgende punten. Vooreerst wordt de realiteitswaarde nagegaan, door o.m. de vraag te beantwoorden of het door Hayes en Flower uitgewerkte expertmodel correspondeert met de reële schrijfprocessen. Vervolgens wordt ingegaan op de algemeengeldigheid. Hiermee wordt bedoeld een uitspraak te formuleren over de toepasbaarheid van het model op een variëteit van schrijfopdrachten. Voorts wordt de overstap van descriptie naar prescriptie aan de orde gesteld. Kan m.a.w. het descriptief model rechtstreeks worden omgezet in prescripties voor een adequate schrijfomgeving. Ten slotte wordt kort ingegaan op het criterium van de helderheid van rapportering.

Paragraaf 4 rondt deze bijdrage bij wijze van besluit af.

1 *Het schrijfmodel van Hayes en Flower*

In deze paragraaf worden de uitgangspunten van het werk van Hayes en Flower verhelderd. Vervolgens wordt het model zelf beschreven als een basis voor de verdere bespreking. Ten slotte wordt ingegaan op de empirische onderbouwing ervan, meer bepaald aan de hand van het probleemoplossingsmodel van Newell en Simon (1972).

1.1 *Uitgangspunten*

Het principale doel van Hayes en Flower was de ontwikkeling van een schrijfdidactiek, vertrekkend van een cognitief-psychologische analyse van schrijfprocessen (Flower & Hayes, 1977; Flower, 1981). Zodoende wilden deze auteurs een bijdrage leveren aan de theorievorming over het schrijven, o.m. door eerst een degelijk schrijfmodel te ontwerpen. In hun empirisch onderzoek is gewerkt met een vergelijkende analyse van het schrijfgedrag bij experten en beginners. De resultaten van hun speurwerk zijn weergegeven in een expertmodel dat duidelijk de kenmerken vertoont van het vroegere werk van Newell en Simon (1972). Het lineaire driefasige schrijfmodel van Rohman en Wlecke (1964), met het strikte onderscheid tussen de voorbereiding ('prewriting'), het uitschrijven ('writing') en het reviseren ('rewriting') wordt verworpen. Bovendien wordt in de lijn van Emig (1971), vooral aandacht geschonken aan de procesanalyse, en dit in tegenstelling met de gangbare produktanalyses.

Het werken met een expertmodel weerspiegelt een aantal veronderstellingen. Vooreerst wordt ervan uitgegaan dat het schrijfproces rationeel verloopt, doorzichtig is en derhalve vatbaar voor analyse. De keuze voor een contrasterende analyse valt samen met een tweede veronderstelling: uit de confrontatie van schrijfproblemen van beginners met het schrijfgedrag van experten zullen adequate schrijfstrategieën worden afgeleid die ook aan modale, c.q. zwakke schrijvers kunnen worden aangeleerd en die derhalve valide zijn voor een schrijfdidactiek. Ten slotte wordt door de auteurs aangenomen, dat de gebruikelijke indeling van het schrijven in vastliggende successieve fasen niet overeenkomt met het psychologisch functioneren van schrijvers.

1.2 *Het schrijfmodel*

Zoals gezegd, beogen Hayes en Flower bij te dragen tot de uitwerking van een adequate schrijfdidactiek. Ter voorbereiding en ondersteuning hiervan, willen ze zowel de aard als de organisatie van de betreffende deelprocessen in een model onderbrengen (Hayes & Flower, 1980b). De term 'organisatie' verwijst zowel naar het moment en de wijze waarop de diverse deelprocessen worden gestart en stopgezet, als naar hun onderlinge interactie. Het model wordt weergegeven in Figuur 1.

Het schrijfmodel omvat de volgende componenten: 1) de taakomgeving, meer bepaald de schrijfopdracht en de tot dusver geproduceerde tekst, 2) het lange-termijngeheugen, i.c. het kennisbestand van de schrijver, 3) het schrijfproces zelf en 4) de monitor, die het hele proces stuurt. Het model geeft weer voor welk probleem de schrijver wordt geplaatst en hoe bij hem/haar het schrijfproces juist verloopt.

Het probleemkarakter van het schrijven bestaat volgens Hayes en Flower vooral in het opheffen van de spanning tussen een rijkdom aan ideeën en de vormgeving ervan in taal, rekening houdend met een aantal inperkingen en/of eisen van de schrijfopdracht, de schrijfdoelen, het lezerspubliek en de formele regels van de taal (Flower & Hayes, 1980a). Hun analyse van schrijfprocessen bij experten kan verder als volgt beschreven worden.

Een goede schrijver gaat bedachtzaam te

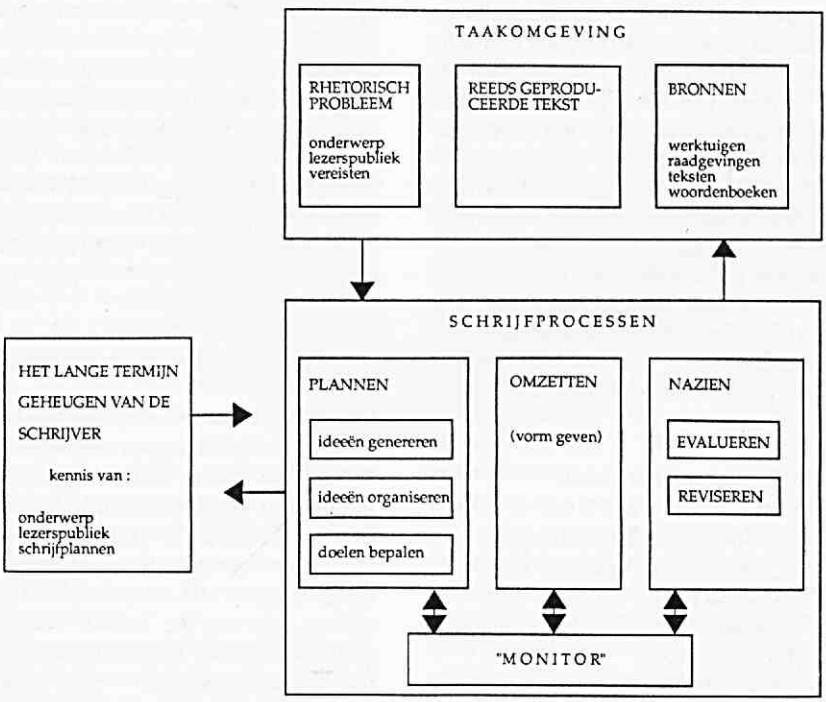
werk. Aandachtig leest hij de schrijfopdracht en analyseert wat van hem precies wordt verwacht. Vervolgens gaat hij over tot de planning, waarin hij nagaat wat hij bij de lezer wil bereiken, welke boodschap overgebracht dient te worden, en met welke gegevens rekening moet worden gehouden. Vanuit een zo volledig mogelijke probleempresentatie, verzamelt de expert ideeën. Hij ordent ze en beslist welke verdere stappen hij nog dient te zetten. Geregeld beoordeelt hij de reeds geschreven tekst door de correspondentie ervan met de opdracht en het schrijfdoel na te gaan. Zo nodig wordt de tekst gereviseerd.

Aldus vallen de volgende fasen te onderscheiden: 1) het plannen ('planning'), 2) het formuleren of in taal omzetten van ideeën ('translating') en 3) het herzien of reviseren, met name: beoordelen en eventueel opnieuw bewerken ('reviewing'). Toch is er een belangrijk verschil met de gangbare lineaire fasering (Rohman & Wlecke, 1964).Een schrijver reviseert de tekst immers niet alleen aan het einde, maar ook tijdens het schrijven zelf. Voor het reviseren tijdens het proces wordt de term 'editing' gehanteerd, terwijl 'reviewing' staat voor de beoordeling en revisie achteraf.

De fasen verlopen derhalve niet in een strikte volgorde, maar ze komen recurrent terug, zij het in een steeds meer geperfectioneerde vorm. Tijdens het formuleren bijvoorbeeld (fase 2) moet de schrijver vaak opnieuw plannen en eventueel zijn ideeën herstructureren (fase 1), ofwel omdat hij door het schrijven nieuwe mogelijkheden heeft ontdekt dan wel omdat het resultaat hem niet (meer) bevredigt. Ook het schrijfdoel wordt tijdens de schrijfact nog herhaaldelijk gewijzigd, daar de schrijver zijn (nog onvoltooide) tekst voortdurend reviseert. Kortom: de subprocessen 'planning', waaronder het genereren van (nieuwe) ideeën, en 'editing' komen in elke fase voor en interfereren met de overige fasen, zij het vooral met het formuleren.

1.3 *Empirische onderbouw van het model*

Wat de empirische ondersteuning van het model betreft, kunnen vragen worden gesteld over de totstandkoming van het model, en over de kwestie of het een resultaat is van empirisch onderzoek dan wel op een reeds bestaand referentiekader voortbouwt. Voor het formuleren van een antwoord wordt vooral een beroep gedaan op het oorspronkelijke werk van Hayes en Flower (1980a, b).



Figuur 1 *Cognitieve processen van het schrijfmodel (naar Flower & Hayes, 1981)*

1.3.1 Onderzoeksopzet

De proefgroep bestond zowel uit experten (universiteitsdocenten en journalisten) als uit beginners (scholieren uit de hoogste klassen van de 'High School'). Hun aantal noch de concrete schrijfopdracht wordt echter vermeld. Uit tal van voorbeelden uit hun publikaties (o.m. Hayes & Flower, 1980a; Flower & Hayes, 1980a, b, 1981) evenals uit de gerapporteerde protocollen kan echter worden afgeleid dat er geen artistiek-literaire produktie wordt vereist, maar wel een expositie, c.q. argumentatie over een of ander onderwerp, zoals het eigen beroep of de vrouwenemancipatie. In de voorbeelden wordt het beoogde lezerspubliek doorgaans wel aangegeven.

Wat de duur van de schrijfactiviteit betreft, ging het telkens om een sessie van ongeveer één uur waarin een tekst diende te worden gemaakt en daarbij hardop werd gedacht. Het aldus verkregen materiaal werd voor een nadere inspectie in protocollen omgezet.

1.3.2 Verwerking en interpretatie

De onderzoeksuitkomsten zijn verkregen door een systematische analyse van het volgend materiaal: 1) de protocollen met de zelfrapportering vanwege de proefpersonen tijdens het schrijven, 2) hun notities, aan de hand waarvan de genese van de tekst kon worden achterhaald, en 3) het eindprodukt. De tijd per eenheid van het schrijfproces werd nauwkeurig geregistreerd. In hun verslaggeving lichten Hayes en Flower (1980b) uitvoerig de werkwijze en onderzoeksresultaten toe aan de hand van een exemplarisch schrijfprotocol. Hieruit kan echter de hoeveelheid door hen bewerkte protocollen niet worden afgeleid noch het aantal proefpersonen voor wie de resultaten gelden. Voor een meer uitvoerige beschrijving van de verrichte protocolanalyse verwijzen we naar Vanmaele en Lowyck (1990).

1.4 *Achtergrond van het model*

Het werk van Hayes en Flower is vooral geïnspireerd door het schrijfonderzoek van Emig (1971) en door de cognitieve psychologie. Welk concreet probleemoplossingsmodel de bron is voor hun schrijfmodel expliciteren Hayes en Flower niet.

Een door ons verrichte analyse toont aan dat vooral het werk van Newell en Simon (1972) het uitgangspunt is geweest. Ten eerste wordt hier het meest naar verwezen (Flower & Hayes, 1977; Flower, 1981; Hayes & Flower, 1980a, b), terwijl andere mogelijke bronnen ofwel niet worden vermeld (Bahle, 1939, zie: Rasch, e.a., s.a.; De Groot, 1961; Duncker, 1935; Selz, 1922) dan wel slechts sporadisch worden genoemd (o.m. Miller, Galanter & Pribram, 1960; Norman & Lindsay, 1975; Norman & Rumelhart, 1975; Rumelhart, 1975). Ons vermoeden van een sterke beinvloeding door Newell en Simon wordt nog versterkt door sommige publikaties die wijzen op een samenwerking tussen Simon en Hayes (Simon & Hayes, 1974, 1976) en op de vertaling door Hayes (1978) van Newell en Simons ideeën. Voorts zijn heel wat begrippen, eigen aan het werk van Newell en Simon, in het schrijfmodel opgenomen, en wel in de exacte verwoording van Newell en Simon. De inhoudelijke vulling ervan is evenwel aangepast aan de eigen aard van de schrijfact. Het gaat om termen zoals: 'taakomgeving' en 'langetermijngeheugen', terwijl processen, zoals 'planning' en 'probleemrepresentatie' in het model een cruciale plaats innemen. Frappant is eveneens de overeenstemming in de uitgangspunten, met name het reduceren van de hele cognitieve activiteit tot het probleemoplossen. Hayes en Flower definiëren het schrijfproces meer bepaald als 1) doelgericht, 2) serieel, 3) hiërarchisch opgebouwd, 4) recursief en 5) door feedback gecontroleerd. De sub 5) genoemde feedback (zie ook: De Groot, 1961; Miller, Galanter & Pribram, 1960; Newell & Simon, 1972; Wiener, 1948) betekent bij Hayes en Flower dat een vrij continue evaluatie van het schrijfproces en van de tot dusver voltooide tekst is ingebouwd.

2 *Een aanvullend referentiekader*

De bovenstaande beschrijving roept een tweevoudige vraagstelling op, met name in hoeverre schrijven wel een vorm van probleemoplossen is en of het in zijn geheel daartoe herleid kan worden. Een afdoend antwoord is slechts mogelijk na een grondige analyse van het probleemoplossen zelf, aangevuld met een aantal andere, niet op het terrein van probleemoplossen liggende benaderingen die mogelijk relevant zijn, zoals kan blijken uit de modellen voor artistieke en sociale processen, de probleemvinding en de schematheorieën.

2.1 *Probleemoplossen als referentiekader*
Om na te gaan in hoeverre het schrijven ook het probleemoplossen veronderstelt, worden eerst de kenmerken van het probleemoplossen en een problementypologie gepresenteerd. Naderhand wordt het schrijven als activiteit hiermee geconfronteerd.

2.1.1 Probleemoplossen
Kenmerken
Een belangrijk kenmerk van het probleemoplossen is dat op de weg van de initiële situatie (de gegevens uit de opgave) naar het doel (de gevraagde oplossing) talrijke hindernissen voorkomen, die slechts vermeden kunnen worden door het inschakelen van hogere denkprocessen (Andriessen & Boonman 1988; De Groot, 1961; Duncker, 1935; Frijda & Elshout, 1978; Newell & Simon, 1972; Voss, Greene, Post & Penner, 1983).

Als voorwaarde daartoe geldt, althans volgens de 'problem solving'-literatuur, de aanwezigheid van hogere cognitieve vaardigheden en van algemene en domeinspecifieke kennis. Hogere cognitieve vaardigheden worden opgedeeld in probleemoplossings- en metacognitieve vaardigheden. De eerstgenoemde zijn de operaties, nodig om de weg naar de oplossing te vinden. Metacognitieve vaardigheden verwijzen zowel naar de kennis over het algemeen met inbegrip van het specifiek individueel cognitief functioneren als naar de bekwaamheid om het eigen denkproces te sturen (Flavell, 1979; Sternberg, 1985). Zo komt men aan de weet welke strategieën geschikt zijn voor de aanpak van een probleem of deelprobleem. Deze kennis echter is slechts effectief zo men ze ook kan aanwenden om het eigen denkproces te sturen, rekening houdend met o.m. de taak (representatie), de eigen doelen en de feedback.

Hogere denkoperaties worden uiteraard uitgevoerd op inhouden, die enerzijds een brede werkelijkheid weerspiegelen en anderzijds in een specifieke discipline kaderen. Vandaar de behoefte aan algemene en domeinspecifieke kennis, naast de vaardigheden in het probleemoplossen.

Problementypologie
De vraag rijst of het model van Newell en Simon zonder meer overdraagbaar is op het omvattende schrijfproces. Een dergelijke extrapolatie vertrekt immers van de veronderstelling dat iedere vorm van probleemoplossen eenzelfde opgavetype bevat en juist daarom gelijkaardige denkoperaties impliceert. Een nadere analyse van vroeger en recent onderzoek over probleemoplossen wijst echter op het tegendeel. Vergelijken we daartoe de opgaven van Selz (1922), Claparède (1933), Duncker (1935), Newell en Simon (1972), Hayes en Flower (1980), Voss, Greene, Post en Penner (1983) met elkaar. De taken van Newell en Simon en de theoretische problemen van Duncker bevatten overwegend symbolisch, die van Claparède vooral figuratief en de overige hoofdzakelijk verbaal materiaal. Maar ook binnen de verbale categorie variëren de vereiste denkoperaties.

Vermits probleemoplossingsmodellen doorgaans op een specifiek opgavetype steunen, vormen ze geen algemeengeldend referentiekader voor de theorievorming over probleemoplossen noch voor het exact omschrijven van het schrijfproces. Een eerste aanzet voor een deugdelijke theorievorming is, aldus Greeno (1978), het uitwerken van een problementypologie.

Voor een classificatie van problemen kunnen twee soorten criteria gelden: enerzijds het type van denkoperaties en onderliggende vaardigheden voor het vinden van de oplossing (Greeno, 1978) en anderzijds de interpretatieruimte (Getzels & Csikszentmihalyi, 1976; Hayes, 1978; Simon, 1973). Op grond van het eerste criterium onderscheidt Greeno drie typen: 1) structuurproblemen ('inducing structure problems'), waarbij de onderliggende structuur moet worden afgeleid uit de gegevens, 2) transformatie- en 3) arrangementsproblemen ('transformation' en 'arranging problems'). Bij het tweede type moet het probleem eerst meerdere malen worden geherformuleerd in termen van een middeldoelstructuur (Duncker, 1935; Frijda & Elshout, 1978; Newell & Simon, 1972), terwijl het derde type het herstructureren betekent van reeds gegeven elementen. De meeste opdrachten zijn een combinatie van minstens twee opgavetypen.

Inzake interpretatieruimte wordt een onderscheid gemaakt tussen precies en vaag omschreven problemen. Hier wordt echter een verschillende terminologie gehanteerd. Simon maakt een onderscheid tussen 'well-structured' en 'ill-structured problems', terwijl Hayes (1978) spreekt over 'well-defined' versus 'ill-defined problems' en Getzels en

Csikszentmihalyi (1976) de 'presented problems' tegenover 'discovered problems' stellen. In een nauwkeurig omschreven probleem bevat de opgave eenduidige gegevens en wordt de vraag in vrij specifieke termen aangegeven: een eerder gesloten taak dus. Zo'n probleem leent zich voor een nauwgezette analyse, met een vrij precieze omschrijving zowel van de criteria waaraan de oplossing moet beantwoorden, als van minstens één wenselijke oplossingsmethode. Van een vaag omschreven probleem is het doel slechts in algemene termen omschreven. De verdere concretisering ligt bij de probleemoplosser en krijgt slechts vorm tijdens het probleemoplossingsproces zelf. Strikt noodzakelijk is het geleidelijk concretiseren van het doel in precieze subdoelen en van het globale probleem in afgebakende subproblemen. Uiteraard betreft het hier geen dichotomie, maar een continuüm.

2.1.2 Schrijven als probleemoplossen

Schrijven en de kenmerken van probleemoplossen

Vooraleer het schrijven te toetsen aan het probleemoplossen, wordt eerst de schrijfactiviteit doorgelicht. Deze kan worden gedefinieerd als het construeren in taal van een voor de schrijver betekenisvolle realiteit, die in een bepaalde vorm is opgeslagen in het lange-termijngeheugen. De keuze en ordening van de kennisaspecten, evenals de verbale vormgeving moeten dan worden afgestemd op het schrijfdoel en het lezerspubliek. Dit veronderstelt een aantal niet als sequentieel op te vatten operaties, m.n.: 1) het analyseren van de (gegeven of zelfgekozen) opdracht, resulterend in een zo accuraat mogelijke representatie ervan, 2) het bepalen van schrijfdoelen, i.c. de beoogde veranderingen bij de lezer, 3) het ontwerpen van een globaal concept van de tekst, 4) het oproepen en representeren van de in het lange-termijngeheugen aanwezige kennisstructuur, 5) het selecteren en herstructureren van concepten en hun relaties, in functie van de schrijfopdracht en van de daarmee verbonden eisen, en 6) de vormgeving van het geheel in taal, met een bijzondere aandacht voor de globale en lokale coherentie (zie o.a. Andriessen & Boonman, 1988; Kintsch & Van Dijk, 1978; McCutchen, 1986).

Het voorgaande wijst erop dat schrijven (althans bij experten) de essentiële kenmerken van het probleemoplossen bevat. Voor het realiseren van het schrijfdoel wordt immers een complex stel hogere cognitieve operaties verondersteld. Bovendien wordt het probleemkarakter nog groter, omdat het lezerspubliek en het beoogde doel de keuzeruimte beperken.

Schrijven en de probleemtypologie

Belangrijk is de constatering dat het model van Newell en Simon en van Hayes en Flower beide gebaseerd zijn op een verschillend opgavetype. De opgaven van Newell en Simon zijn transformatieproblemen met een gesloten karakter. Bovendien worden ze overwegend in de vorm van symbolisch materiaal aangeboden, in tegenstelling met het schrijven, dat het omgaan met verbaal materiaal veronderstelt.

Vertrekkend van de genoemde typologie kunnen de volgende specifieke kenmerken van het schrijven als probleemoplossen nader worden geïdentificeerd, zij het niet exhaustief.

1. Volgens de definitie van schrijven dient een vrij groot aantal specifieke operaties wellicht opgevat te worden als de combinatie van een transformatie-en een arrangementsprobleem (Greeno, 1978).
2. Schrijfopdrachten vertonen kenmerken van een vaag omschreven probleem (Simon, 1973; Hayes, 1978; Getzels & Csikszentmihalyi, 1976). Uiteraard kan ook hier de interpretatieruimte veriëren naargelang van de vrijheidsgraden in de opdracht en van de interpretatie van de opdracht door de schrijver (Getzels & Csikszentmihalyi, 1976). Het open karakter impliceert alleszins 'problem finding'.
3. Bij schrijven is het probleemoplossingsproces weinig doorzichtig en het verloopt zeker niet lineair.
4. Naast probleemoplossen op zich, impliceert schrijven het verbaal representeren van een probleemoplossingsproces dat voor de lezer zo doorzichtig mogelijk moet zijn. Daar komt nog bij, dat het een diversiteit van redeneringen veronderstelt, o.m. om de stappen in het probleemoplossingsproces te ondersteunen. Dit laatste geldt vooral voor expositorische teksten.

2.2 *Aanvullende theoretische invalshoeken*

De specificiteit van de schrijfactiviteit roept de behoefte op van een exploratie van moge-

lijke aanvullende kaders. Daarom wordt eerst aandacht besteed aan modellen die vrij nauw kunnen aansluiten bij het schrijfproces, zoals: een model voor artistieke processen, en een redeneermodel voor sociale processen. Aangezien in een aantal ervan de 'probleemvinding' een rol speelt, wordt dit aspect verder meer uitvoerig toegelicht. Ten slotte wordt aandacht besteed aan de betekenis van de schematheorieën voor het schrijven.

2.2.1 Model voor artistieke processen

Reeds in het verleden werden aspecten van het probleemoplossen teruggevonden in de studies over artistieke processen. Bahles procesmodel voor muzikale composities (1939) bouwt bijvoorbeeld verder op het werk van Selz (1922), die één probleemtype grondig beschrijft, met name het analyseren van bestaanservaringen om er vorm aan te geven in een artistieke creatie. Uitgaande van de gangbare definitie van probleemoplossen (Duncker, 1935; Newell & Simon, 1972), ontwerpen Getzels en Csikszentmihalyi (1976) eveneens een procesmodel voor de beeldende kunsten. In die modellen worden drie gemeenschappelijke processen onderscheiden: 1) de probleemformulering bestaande in het ontwerpen van het globale concept van het kunstwerk, zoals de thematiek, 2) het probleemoplossen, m.n. de vormgeving ervan in klanken of beelden, en 3) het evalueren. Deze drie processen corresponderen ten dele met de planning, de omzetting en het reviseren in het model Hayes en Flower.

2.2.2 Redeneermodel voor sociale processen

Met het werk van Newell & Simon (1972) als vertrekpunt, ontwikkelden Voss e.a. (1983) een model dat afgestemd is op de specifieke eigenschappen van sociale problemen. Kenmerkend is het belang dat wordt gehecht aan de verantwoording van de diverse stappen in het probleemoplossingsproces, wat dan weer een stel redeneeroperaties insluit. Om die reden wordt door de auteurs een redeneringsstructuur in de probleemoplossing ingebouwd. Meer bepaald worden niet alleen de stappen in het probleemoplossen gespecificeerd, maar evenzeer de opeenvolgende redeneringen voor de verantwoording van de genomen beslissingen.

2.2.3 De 'problem-finding'

In recente probleemoplossingsmodellen, o.m. in dit van Newell en Simon, komt de probleemvinding ('problem-finding') zelden ter sprake. Voor meer informatie omtrent deze kwestie zijn we dus aangewezen op oudere, meer algemene denkmodellen (zie: Dewey, 1910; Claparède, 1933) of op de reeds genoemde beschrijving door Bahle (1936, 1939) van artistieke processen (zie: Rasch, s.a.; Getzels & Csikszentmihalyi, 1976). Dit hoeft geen verbazing te wekken, aangezien deze modellen steunen op een open opgaventype. In tegenstelling met Brugman (1989) beschouwen we het model van Getzels en Csikszentmihalyi wel degelijk als een model van probleemvinding. Vooreerst bood de opdracht vele kansen tot probleemvinding. De studenten werden niet zomaar verzocht om, op grond van een keuze uit zevenentwintig objecten, een stilleven te tekenen. Wel fungeerden de gepresenteerde voorwerpen als een stimulans voor de exploratie van een persoonlijke realiteitservaring. Ten tweede wordt de probleemvinding voldoende nauwkeurig omschreven en geoperationaliseerd in het onderzoeksopzet, o.m. in de aangeboden probleemsituatie, de dataverzameling en -verwerking (zie ook: Vanmaele & Lowyck, 1990).

In het proces van de probleemvinding kan men twee aspecten onderscheiden: 1) het zelf construeren van de opgave of van het probleem, door o.m. zelf het probleem (doel, thema, lezerspubliek) te bepalen (zie: Bahle, 1939; Rasch, s.a. ; Getzels & Csikszentmihalyi, 1976), en 2) het preciezer omschrijven van het probleem. Dit laatste kan betekenen: het identificeren van de knelpunten die de kloof tussen de gegevens en de oplossing veroorzaken (Dewey, 1910), het in een opgave ontdekken en expliciteren van het gevraagde (als resultaat van de doelanalyse) en het geleidelijk duidelijker formuleren van een hypothese (Claparède, 1933).

De juiste plaats van de 'problem-finding' in het hele probleemoplossingsproces wordt door sommigen opgevat als de eerste fase in het probleemoplossen (Bahle, 1939; Dewey, 1910), terwijl anderen het spreiden over het gehele probleemoplosproces (Claparède, 1933; Getzels & Csikszentmihalyi, 1976). Voor een verdere analyse van de probleemvinding verwijzen we naar Brugman (1989).

2.2.4 De schematheorieën

De bovenstaande modellen representeren weliswaar de cognitieve processen, maar ze zeggen weinig of niets over de kennisstructuur van de probleemoplosser, de wijzigingen die het oplossen van een probleem daarin teweeg brengt, e.d. Voor deze kwesties kan men beter terecht bij de schematheoretische benadering (Kintsch & Van Dijk, 1978; Norman & Rumelhart, 1975; Schank, 1975).

Schematheorieën zijn voor het bestuderen van het schrijven waardevol, vooral wat betreft de exploratie van vragen naar de inhoud en de organisatie van het kennisbestand van de schrijvende leerling of naar de invloed hiervan op schrijfproces en -produkt. Ze bieden bovendien een bruikbaar referentiekader voor de aanmaak van criteria voor de inhoudelijke en structurele analyse van schrijfprodukten (zie: Kintsch, 1979; Kintsch & Van Dijk, 1978).

Dit theoretisch kader is niet alleen relevant voor expositorische teksten, maar evenzeer voor artistieke of ludieke produkties. Zo wordt bij het schrijven van een cursiefje in de krant een beroep gedaan op enerzijds de reeds aanwezige kennis van een stukje werkelijkheid, dat op een originele en ludieke wijze wordt ge(re)construeerd, en anderzijds op een conventioneel genreschema. Dit laatste verwijst naar de inhoud, structuur en taalkundige conventies, die specifiek zijn voor een bepaald tekstgenre (zie o.m. Kintsch & Van Dijk, 1978). Daardoor worden voor de diverse tekstgenres ook uiteenlopende schemata benut. In een cursiefje wordt van andere schema's gebruik gemaakt dan in een wetenschappelijk artikel, een brief of een gedicht.

3 *Beoordeling van het model van Hayes en Flower*

Na de beschrijving van het model van Hayes en Flower en van een aantal mogelijke aanvullende invalshoeken zijn we thans aan een beoordeling toe. Zoals is aangetoond, wilden Hayes en Flower de schrijfprocessen van experten representeren in een model, om van daaruit een schrijfdidactiek te ontwikkelen (Flower, 1981).

Die dubbele intentie doet de behoefte ontstaan om aan de hand van een aantal concrete criteria de waarde van hun schrijfmodel te onderzoeken. Men dient er immers van uit te gaan dat het model alle essentiële aspecten van het schrijven bevat, zodat het correspondeert met de werkelijke schrijfprocessen. Het moet bovendien toepasbaar zijn op vele individuen of soorten groepen en op uiteenlopende schrijfsituaties. Daaraan gekoppeld is het criterium van de verantwoorde overgang van een descriptie (i.c. van expertgedrag) naar prescriptieve regels voor beginnende schrijvers. Ten slotte geldt het criterium van de helderheid en eenduidigheid bij de rapportering ten behoeve van een eventuele controle en replicatie van het verrichte onderzoek.

Het schrijfmodel van Hayes en Flower wordt thans aan de vier genoemde criteria getoetst.

3.1 *Realiteitswaarde*

Hier komt de vraag aan de orde of en in hoeverre het door Hayes en Flower uitgewerkte expertmodel correspondeert met de reële schrijfprocessen en of het als raamwerk voldoet voor de analyse van het schrijven als een omvattende activiteit. Er zal daarom worden nagegaan welke eigenschappen van het schrijven in het voorgestelde model voorkomen.

3.1.1 De eigenschappen van de schrijfact

De meeste modellen die eerder in dit artikel zijn besproken bevatten kenmerken, die ongetwijfeld ook tot het schrijven behoren. Gedacht wordt aan de redenering, de probleemvinding, en aan het dynamische, niet-lineaire karakter van het schrijfproces. Hierop wordt kort ingegaan.

Omwille van de concrete beschrijving van de *redeneeroperaties* voor de ondersteuning van het oplossen van sociale problemen, biedt het model van Vöss e.a. (1983) tal van aanknopingspunten met het schrijven. Alhoewel Hayes en Flower in 1980 niet naar dit meer recente model konden verwijzen, lijken de kenmerken ervan voor het schrijven erg interessant. De redeneerstappen kunnen met name als een vorm van 'coaching' fungeren, althans voor expositorische teksten waarin de redenering juist een centrale rol krijgt toebedeeld. Hetzelfde geldt voor de modellen betreffende de artistieke processen (Bahle, 1936; Getzels & Csikszentmihalyi, 1976), omdat ook hierin zowel de probleemvinding als het dynamische karakter van het schrijven tot hun recht komen.

De *probleemvinding* lijkt alvast een zeer belangrijk aspect van het schrijfproces, omdat de schrijfopdracht dikwijls een open opgave is, zodat de schrijver het doel en de daarmee samenhangende restricties zelf moet vastleggen. Schrijven is dus geen loutere probleem-'oplossing', maar evenzeer een probleem-'detectie'. Hayes en Flower brengen de probleemvinding inderdaad ter sprake, en ze gebruiken hiervoor de uitdrukking 'making meaning' of 'making problems' (Flower & Hayes, 1980b; Hayes, 1978). Immers, de term 'problem-finding', c.q. 'discovery' suggereert volgens hen dat er iets aanwezig is dat alleen nog maar ontdekt dient te worden. Het gaat echter niet om het vinden van een reeds bestaand, maar wel om het construeren van een nieuw probleem. Meer bepaald vatten Flower en Hayes (1980b) probleemvinding op als het representeren van een – i.c. retorisch – probleem. In de lijn van Getzels en Csikszentmihalyi (1976) beschouwen ze de probleemvinding niet als een (eerste) fase van een probleemoplossingsproces. Het is wel verbonden met een specifiek probleemtype, m.n. een vaag omschreven probleem, en het maakt dan ook een inherent deel uit van het hele probleemoplossingsproces. Een vaag omschreven probleem leent zich immers niet voor een volledige representatie vooraf. De probleemoplosser, i.c. de schrijver moet zich integendeel 'in het probleem storten', weliswaar vanuit een aanvankelijke en partiële probleemrepresentatie (Hayes, 1978). Doelbepaling en planning vinden (althans bij een expert) wel uitdrukkelijk gedurende de eerste fase plaats, maar ze zijn voortdurend aan verandering onderhevig. De progressieve probleemverkenning levert vervolgens geleidelijk aan meer informatie op, en dit vereist een nagenoeg voortdurende herformulering van de doelen en plannen.

De voorgaande omschrijving suggereert een volledige integratie van 'problem-finding' in het model. Toch hebben we de indruk dat het er geen wezenlijk deel van uitmaakt, maar er eerder is aan toegevoegd. Probleemvinding komt slechts in een beperkt aantal publikaties voor (Flower & Hayes, 1980b) en het wordt bovendien niet opgenomen in de beschrijving van het model als geheel. Dit laatste geldt wel voor de reeds genoemde modellen van Bahle, Claparède, Dewey, en van Getzels en Csikszentmihalyi. De vraag is dan ook of het door Hayes en Flower gebruikte concept 'probleemrepresentatie' de eigenheid van de probleemvinding wel voldoende weergeeft, of – in andere woorden – 'problem-finding' wel helemaal te herleiden valt tot een adequate probleemrepresentatie.

Een ander specifiek kenmerk van schrijven, en a fortiori van elk artistiek en creatief proces, is zijn *dynamisch en niet-lineair karakter*. In hun concrete omschrijving van het model worden zowel de recursiviteit (Hayes & Flower, 1980a) als het dynamisch karakter van het schrijfproces (Flower & Hayes, 1980b) benadrukt. Maar de grafische weergave (Figuur 1) representeert een eerder statische opvatting: een soort variant van het driefasenmodel van Rohman en Wlecke (1964). De nadere inspectie van Hayes en Flowers publikaties wijst overigens op een toenemende discrepantie tussen de verbale omschrijving van hun steeds dynamischer wordende opvatting over het schrijven en de statische modelmatige weergave ervan. Daaruit blijkt onzes inziens de beperktheid van een omzetting van de essentie van een fenomeen in een visueel model. Hetzelfde geldt trouwens evenzeer voor de modelrepresentaties in het werk van Claparède (1933), Bahle (1936), en Getzels en Csikszentmihalyi (1976).

3.1.2 Schrijven als omvattende activiteit

Het is reeds gezegd, dat het door Hayes en Flower uitgewerkte model niet de heelheid van het schrijven weergeeft. Dit is trouwens van toepassing op iedere modelmatige benadering. Toch willen we wijzen op een terzake belangrijke tweevoudige reductie.

Een eerste inperking, overigens inherent aan de meeste probleemvindings- en probleemoplossingsmodellen, is de mentalistische opvatting van het proces, i.c. het schrijven. De analyse van de cognitieve processen staat er centraal, zodat weinig aandacht uitgaat naar aspecten als: de inhoud, de kennisstructuur in het lange-termijngeheugen, de interactie ervan met de schrijfvaardigheid (Eigler, Jechle, Merzinger & Winter, 1987), de eisen aan de tekststructuur (Kintsch & Van Dijk, 1978), en de rol van de schrijfopdracht. Hoogstens wordt in Flower en Hayes (1984) ingegaan op de structuur van de reeds aanwezige kennis in het lange-termijngeheugen (Bartlett, 1932; Norman & Rumelhart, 1975; Schank, 1972), en op de representatie ervan in

taal. Dit mentalisme komt eveneens tot uiting in de onderzoeksopzet. Hayes en Flower beperken zich met name tot een procesanalyse, terwijl de analyse van de schrijfprodukten zelf wordt nagelaten. Een integratie van de schematheorieën zou alvast bijdragen tot een verrijking van het schrijfonderzoek.

Maar er is meer: niet alleen wordt het schrijven als cognitieve activiteit gereduceerd voorgesteld, maar het wordt ook helemaal tot een cognitieve activiteit herleid. Dit komt duidelijk naar voren in het ontbreken van de hierna volgende aspecten. Vooreerst worden dynamisch-affectieve aspecten buiten beschouwing gelaten. 'Doelgerichtheid' (het dynamische aspect) verwijst bij Hayes en Flower alleen naar het (bewust of onbewust) nastreven van min of meer concrete schrijfdoelen. Schrijven wordt er niet in verband gebracht met de bredere motieven van de schrijver, en evenmin wordt gepeild naar de persoonlijke betekenis van een schrijfactiviteit voor het individu of naar de oorsprong van diens schrijfdoelen.

In de oudere denkpsychologie werd probleemoplossingsgedrag wel beschouwd als een motivationele en cognitieve eenheid. Een ontkoppeling merken we reeds bij Miller, Galanter en Pribam (1960), die bijvoorbeeld 'intentie' omschrijven als "gestart zijn met de uitvoering van een plan" en niet meer verwijzen naar de dynamische onderbouw ervan (Kuhl, 1983). Het affectieve krijgt evenmin aandacht, zoals o.m. de invloed van enthousiasme, onzekerheid en beleving van een schrijfsituatie op het schrijfproces en -produkt. Ten tweede ontbreekt de aandacht voor de bredere context. Er is alleen sprake van een taakomgeving (schrijfopdracht, hulpmiddelen en de tot dusver geproduceerde tekst). Welke invloed gaat echter uit van persoonlijke normen en criteria die van het moedertaalonderwijs zijn overgeërfd, of van de taalcultuur thuis? In welke andere opzichten zijn het moedertaalonderwijs, de groepsstructuur in vormings-. c.q. arbeidsmiddens en het gezin als de zovele socio-culturele gegevens bepalend voor het schrijven?

Deze reductie is overigens kenmerkend voor de hele problem-solving benadering. Opmerkelijk is echter dat anderen, zoals Newell en Simon (1972) de eliminatie van dynamisch-affectieve aspecten expliciet verantwoorden. Volgens hen is het speurwerk ter zake nog onvoldoende gevorderd en moet eerst het onderzoek van cognitieve processen op punt worden gesteld. Ook al is deze stelling voor discussie vatbaar, toch verdient de explicitering van genoemde beperking waardering. Ons bezwaar kan derhalve niet slaan op een reductie als dusdanig. De focusering op slechts één dimensie is immers vaak aangewezen, ook al zijn ons meer holistische onderzoeksdesigns bekend (Rose, 1985). Bezwaarlijk is wel dat de gekozen beperking door Hayes en Flower niet wordt geëxpliciteerd en dat een situering van de gekozen modelcomponenten in een ruimer kader ontbreekt.

3.2 *Algemeengeldigheid*

Of Hayes en Flower hun model als een algemeengeldend expertmodel voorstellen is niet duidelijk uit hun publikaties af te leiden. O.i. kan het alleszins niet als zodanig worden aangediend. Uit het voorgaande blijkt immers reeds dat het model niet alle belangrijke componenten van het schrijven als een vorm van probleemoplossen omvat. Bovendien gelden hun bevindingen wellicht niet voor een diversiteit van schrijfsituaties. Met name werd bij het onderzoek geen variëteit in de schrijfopdrachten ingevoerd. Artistiek-literaire prestaties, bijvoorbeeld, werden uitgesloten, misschien omwille van de overwegend cognitieve benadering. Bovendien is de protocolanalyse gebaseerd op één schrijfsessie van ongeveer een uur per proefpersoon. Een eenmalige schrijfsessie biedt nog weinig zicht op iemands schrijfpatroon. Ten tweede verschaffen sessies van een beperkte duur nagenoeg geen informatie over de langetermijnplanning, die men toch van een goed schrijver mag verwachten. Het belang van de duur werd ook in een door ons verrichte voorstudie geconstateerd (zie: Vanmaele, 1990). Schriftelijke taalvaardige leerlingen vervielen in de 'knowledge telling strategy', wanneer de schrijfopdracht slechts in een sessie van vijftig minuten werd uitgewerkt. Maar bij een schrijfopdracht over een week gespreid, werd veel meer inhoudelijk gepland. Volgens de proefpersonen werkt tijdsdruk erg remmend en dit staat de geleidelijke groei van een tekst in de weg.

Niet alleen de duur, maar ook de aard van de schrijfopdracht, i.c. de gepercipieerde moeilijkheidsgraad ervan, de tekstlengte en het tekstgenre, beïnvloeden de schrijfstrate-

gieën en de planning (Molitor, 1987). Tot nog toe werd opgemerkt dat de onderzoeksbevindingen van Hayes en Flower wellicht niet voor alle schrijfsituaties (o.m. opdrachten, duur, context) gelden. Een verdere vraag is of ze veralgemeenbaar naar diverse individuen toe zijn. Persoonsspecifieke verschillen zijn namelijk evenmin in het onderzoek van Hayes en Flower ingebouwd. Nochtans spelen de kenmerken van de schrijver, zoals o.m. de vertrouwdheid met het tekstgenre en het aantal beschikbare tekstschemata, een rol (Molitor, 1987). Bovendien volgt het schrijven niet voor alle individuen een neerwaartse ('top-bottom') beweging, wat Hayes en Flower wel stellen. In dit opzicht trof Molitor (1987) met name twee soorten experten aan. De eerste categorie werkt inderdaad op een neerwaartse wijze: vanuit een zo volledig mogelijke probleemrepresentatie bepalen deze personen hun schrijfdoelen en maken ze vooraf de uitgewerkte planning. Andere experten daarentegen gaan vrij spoedig over tot het formuleren van hun tekst. De schrijfdoelen, opties en hun ideeën krijgen slechts vorm in interactie met de groeiende tekst, wat wijst op een opwaartse ('bottom-up') beweging. Eerstgenoemde schrijvers prepareren een groot aantal notities vooraf, terwijl de anderen hun tekst herhaaldelijk reviseren.

Onderzoek aan de hand van een diversiteit van schrijfopdrachten en -condities is, naast de aandacht voor persoonsspecifieke verschillen eveneens aangewezen. In dit verband bouwden wij in ons onderzoek een – zij het uiterst beperkte – diversiteit van schrijfopdrachten in, door de leerlingen te confronteren met een verhaal en met een expositorische tekst. Inzake schrijfcondities voorzag Molitor (1987) nog een aanvulling op het model van Hayes en Flower, door langetermijnopdrachten te voorzien. Ditzelfde doen o.m. Emig (1971) en Bahle (1939).

De eis van algemeengeldigheid is weliswaar niet absoluut, aangezien ieder onderzoek beperkingen heeft inzake de steekproeftrekking en de soort onderzoekssituaties. Noodzakelijk is echter dat de ingevoerde beperkingen worden geëxpliciteerd en dat wordt afgebakend voor welke individuen en schrijfsituaties het gebruikte model geldt. Juist deze informatie ontbreekt in de rapportering door Hayes en Flower.

3.3 *De overstap van descriptie naar prescriptie*

Op basis van expertonderzoek wordt door Hayes en Flower een aantal voorschriften ten behoeve van een schrijfdidactiek voor de modale leerling ontworpen (Flower, 1981). Daaraan ten grondslag ligt de veronderstelling, dat de modale, c.q. schriftelijk taalzwakke leerlingen expertstrategieën kunnen overnemen, mits voldoende didactisch ondersteund. Nochtans kunnen de didactische interventies niet rechtstreeks vanuit expertonderzoek worden afgeleid, vooral omwille van aanzienlijke verschillen in de schrijfstrategieën van professionele en beginnende schrijvers (Hayes & Flower, 1980a, 1980b; Molitor, 1987; Scardamalia & Bereiter, 1987; Scardamalia & Paris, 1985).

Om uit te kunnen maken of de gevonden expertstrategieën rechtstreeks aan beginners kunnen worden aangeleerd, dient een aantal bruggen te worden geconstrueerd. Vooreerst moeten nauwkeurige hypothesen worden geformuleerd over het reële schrijfgedrag van jonge leerlingen, over de kenmerken ervan en over de veronderstelde effectiviteit van specifieke interventies. Daarop volgt een construerende fase, waarin de geconcipieerde interventies geleidelijk aan worden verfijnd. Dit dient om de gestelde hypothesen nader te toetsen en om te controleren of en in welke mate de interventies wel degelijk aansluiten bij de actuele kenmerken en de ontwikkelingsmogelijkheden van de doelgroep (o.m. wat het concrete schrijfgedrag en de onderliggende processen betreft). De prescripties voor een schrijfdidactiek die aldus zijn ontwikkeld, dienen aansluitend te worden uitgetest bij representatieve individuen en (klas-)groepen, en dit onder gecontroleerde omstandigheden.

Uit het bovenstaande moge blijken, dat Hayes en Flower de complexiteit van de overstap van descriptie naar prescriptie onderschat hebben, aangezien er niet gewerkt is aan de bruggen met behulp van gecontroleerde constructies.

3.4 *Helderheid in de rapportering*

In diverse opzichten worden gangbare eisen inzake helderheid van rapportering niet ingelost. Op het tekort aan helderheid over de oorsprong van hun model is reeds eerder gewezen. Problematisch is verder het gebrek aan precisie in de beschrijving van de onderzoeks-

opzet. Alleen de protocolanalyse wordt nauwgezet beschreven, zij het enigszins beperkend aan de hand van een voorbeeld. Vragen over de omvang van de proefgroep blijven onbeantwoord, en het type van schrijfopdrachten valt slechts uit de – weliswaar goed gestoffeerde – voorbeelden af te leiden.

Criteria voor de keuze en de presentatie van de opdrachten worden evenmin geëxpliciteerd, terwijl de allocatie van een proefpersoon bij de groep experten, c.q. beginners niet wordt verantwoord. Wellicht baseren de auteurs zich hiervoor op de proceskenmerken van het schrijfgedrag en niet op de kwaliteit van het schrijfprodukt. De maat voor een eventuele effectmeting blijft alvast impliciet. Hoogstens vindt men soms criteria die door de proefpersonen zelf worden gehanteerd (o.m. Hayes & Flower, 1980a, b). Vermoedelijk vinden de onderzoekers het vooropstellen van produktcriteria overbodig, daar ze vooral het schrijfproces en niet het schrijfprodukt onderzoeken. De genoemde onduidelijkheid bemoeilijkt in ieder geval zowel de interpretatie als de replicatie van het onderzoek.

Een derde knelpunt is de gebrekkige aflijning van de gebruikte concepten. Vaak moet de betekenis ervan worden afgeleid uit voorbeelden of wordt een concept slechts in een latere publikatie verhelderd. Begrippen als 'strategie', 'procedure', e.a. worden slechts sporadisch gedefinieerd. Hetzelfde geldt voor de concepten 'doel', 'intentie', 'planning' en 'plan'. Voor een nadere inspectie van de conceptuele kwestie zij verwezen naar Vanmaele en Lowyck (1990).

4 *Besluit*

De uitgangsvraag van deze bijdrage was of en in hoeverre het schrijfmodel van Hayes en Flower een voldoende ruim referentiekader biedt voor een omvattende beschrijving van schrijfprocessen, mede met het oog op het ontwerpen van een adequate schrijfomgeving. Na eerst te zijn ingegaan op het model zelf en op een aantal vragen betreffende de relevantie ervan voor het schrijven als activiteit, is het model besproken aan de hand van vier beoordelingscriteria: de werkelijkheidsgetrouwheid, de algemeengeldigheid, de overgang van descriptie naar prescriptie, en de helderheid van de rapportering. Om een gefundeerd oordeel toe te laten, is in de eerste twee paragrafen basisinformatie over het model en zijn achtergronden gepresenteerd.

Bij het onderzoek naar de werkelijkheidswaarde stond het schrijven als probleemoplossen centraal, juist omdat Hayes en Flower hun model als zodanig aandienen. Hierbij vroegen we ons af 1) of en in hoeverre een probleemoplossingsmodel het schrijven als cognitief proces weerspiegelt, m.a.w. of schrijven wel probleemoplossen is, en 2) of schrijven in zijn geheel kan herleid worden tot een cognitief proces.

Wat de eerste vraag betreft is schrijven, althans bij experten, een specifieke vorm van probleemoplossen. Een aantal concepten uit de 'problem-solving'-literatuur, eveneens gebruikt door Hayes en Flower, zijn hierop dan ook van toepassing. Maar een model dat ontworpen werd naar analogie met dit van Newell en Simon (1972) voor het oplossen van mathematische problemen, voldoet niet volledig als referentiekader, zelfs niet indien het schrijven uitsluitend wordt opgevat als een cognitief proces. Voor het verwerpen van deze strikte analogie gelden minstens twee redenen. Ten eerste schiet het begrippenkader tekort voor wat een aantal specifieke kenmerken van schrijfprocessen betreft. Ten tweede weerspiegelt het model een mentalistische opvatting omwille van de eenzijdige aandacht voor (interne) cognitieve processen, met verwaarlozing van de inhoudelijke aspecten (zie: Norman & Rumelhart, 1975; Schank, 1972). M.a.w.: er wordt van uitgegaan dat de formele kenmerken van het schrijven een voldoende basis zijn voor het beschrijven van de omvattende schrijfact. Reeds hoger (zie: par. 2.2.4) is bij de bespreking van de schematheorieën gewezen op het belang van de verschillende inhouden, structuren en taalkundige conventies die juist door een tekstgenre worden bepaald, en waardoor schrijven, zelfs in het geval het om een erg open schrijfopdracht gaat, nooit tot een louter formele activiteit kan worden gereduceerd.

Het specifieke van het schrijven, in vergelijking met andere activiteiten, wordt door Hayes en Flower onderkend. Voor de conceptualisering ervan doen ze ten dele een beroep op algemene concepten uit de 'problem-solving'-literatuur, die ook bij Newell en Simon een belangrijke plaats innemen. Zo wordt, voor wat het dynamisch en niet-

lineaire karakter van het schrijfproces betreft verwezen naar het concept 'recursiviteit' en voor de 'problem-finding' naar de 'probleemrepresentatie'. Andere bronnen, c.q. modellen worden echter slechts sporadisch ter sprake gebracht, en dit voor een beperkt aantal aspecten. Zo verwijzen Hayes en Flower naar het werk van Getzels en Csikszentmihalyi in verband met de plaats van de 'problem finding' in het probleemoplossen. Onzes inziens is hun schrijfmodel alleszins aan verrijking toe door onder meer dieper in te gaan op de 'problem-finding'-aspecten, de redeneringsfuncties, de schematheorieën en de concepten uit de linguïstiek.

De tweede vraag, m.n. in hoeverre het model van Hayes en Flower de totaliteit van de schrijfrealiteit dekt, werd ontkennend beantwoord. Onze bezwaren gelden niet zozeer de rationaliteit van de representatie van het schrijfproces. Pogingen om menselijke handelingen, waaronder het schrijven, doorzichtig te maken, houden uiteraard nog geen ontkenning in van meer 'irrationele' elementen. Problematisch is wel dat het reductionistisch karakter van het schrijfmodel, waardoor het schrijven louter tot een cognitief proces wordt herleid, niet wordt geëxpliciteerd of in de totaliteit van het schrijven gesitueerd.

Wat de algemeengeldigheid betreft (het derde criterium), kan het model niet als zodanig worden aangediend, meer bepaald omdat het onderzoeksmateriaal de diversiteit van schrijfsituaties niet dekt (o.m. inzake opdracht en duur). Bovendien werd weinig of geen rekening gehouden met de mogelijke persoonsspecifieke verschillen in schrijfstrategieën.

Aan het derde beoordelingscriterium, m.n. het waarborgen van een verantwoorde overstap van descriptie naar prescriptie wordt evenmin voldaan. In verband hiermee tekenen we bezwaar aan tegen het rechtstreeks afleiden van voorschriften voor een schrijfdidactiek uit de analyse van expertgedrag (Flower, 1981), zonder enig construerend onderzoek als brug. Door zo te werk te gaan, wordt niet gegarandeerd dat de uitkomsten van het beschrijvend onderzoek reëel geïmplementeerd zijn in de didactische schrijfomgeving. Het tekort aan een systematische controle van iedere stap bij de constructie van leeromgevingen, vormt dan ook het tere punt in veel onderzoeks- en ontwikkelwerk binnen de onderwijskundige ontwerpkunde.

Wat de duidelijkheid bij de rapportering betreft (het vierde criterium), beantwoordt het model slechts in een geringe mate aan de gestelde eisen. Gebrek aan precisie in de rapportering van onderzoeksopzet en -resultaten en in de verheldering van concepten vormen een hinderpaal voor een nauwkeurige interpretatie en replicatie van de onderzoeksuitkomsten.

De genoemde punten van kritiek betekenen echter niet, dat het schrijfmodel van Hayes en Flower als irrelevant mag worden afgedaan. Vooreerst moet het in de tijd worden gesitueerd. Het kwam tot stand in een periode waarin de problem-solving benadering zich vooral beperkte tot het domein van de wiskunde en tot de studie van expertgedrag. Verdienstelijk is dan ook, dat de taal als vakdomein vanuit dit model werd geëxploreerd. Bovendien was de verbreding van de cognitieve psychologie naar een 'cognitieve wetenschap' met een ruime aandacht voor inhoudsspecifieke en linguïstische aspecten (Norman & Rumelhart, 1975; Schank, 1972) nog niet echt doorgedrongen in het cognitief psychologisch onderzoek. Vandaar wellicht het mentalisme van de meeste probleemoplossingsmodellen uit die periode. Positief uitgedrukt betekent het onderzoek van Hayes en Flower een doorbraak in het procesgericht schrijfonderzoek. Het model is hoe dan ook een bruikbaar startpunt voor het identificeren en nader analyseren van de belangrijkste procescomponenten, c.q. fasen van een schrijfproces bij experten en van de daartoe benodigde oplosvaardigheden.

Gezien zijn beperkingen is een exclusieve keuze voor dit schrijfmodel als enig referentiekader evenwel niet verantwoord. Met het oog op schrijfonderzoek bij modale, c.q. taalzwakke leerlingen, en in het perspectief van de constructie van een schrijfdidactiek, is een aanvulling noodzakelijk in minstens twee opzichten.

Ten eerste dient gewerkt aan een specifiek referentiekader voor de diagnose van het schrijfgedrag van modale, c.q. taalzwakke leerlingen. Uit het schrijfmodel kan immers wel worden afgeleid over welke probleemoplossingsvaardigheden jonge schrijvers (nog) niet beschikken, maar niet hoe ze werkelijk schrijven. Hiervoor zijn we aangewezen op het 'knowledge telling model' van Bereiter en Scardamalia (Bereiter, 1980; Scardamalia

& Bereiter, 1987) en op eigen empirisch onderzoek. Door het vatten van de wijze waarop dit soort leerlingen reëel te werk gaan tijdens het schrijven kan alvast een beeld worden geschetst van de eigenschappen van het 'spontane' schrijfproces bij zwakkere leerlingen, wat naderhand geconfronteerd kan worden met het reeds geconstateerde expertgedrag. Ook het 'knowledge transforming model' (Scardamalia & Bereiter, 1987) biedt hier hulp, aangezien het een brug slaat tussen expert- en novietengedrag, o.m. door na te gaan hoe kennis geleidelijk aan getransformeerd wordt op de weg van beginner tot expert en zodoende een helderder licht werpt op de evolutie van het schrijfgedrag van leerlingen.

Ten tweede is nader onderzoek van de meer inhoudelijke aspecten wenselijk, zoals: de invloed van de algemene en domeinspecifieke kennis op het schrijfproces en -produkt, de invloed van schrijven op de ontwikkeling van die beide kennisvormen (Eigler et al., 1987), de kenmerken van schrijfprodukten van modale leerlingen. Voor dit laatste zijn we eerder aangewezen op de schematheorieën en op een linguïstisch referentiekader.

In hoeverre het schrijfmodel van Hayes en Flower een valide basis vormt voor didactische interventies moet uit goed gecontroleerd construerend onderzoek blijken. En, vermits bij het schrijven (ook bij experten) dynamisch-affectieve en sociaal-culturele aspecten eveneens een belangrijke rol spelen, dient voor het schrijfonderzoek een meer integratief raamwerk te worden ontworpen, waarin deze variabelen naast de cognitieve in een schrijfdidactiek zijn opgenomen (Rose, 1985).

Literatuur

Andriessen, J.E.B. & J.H. Boonman, Probleemoplossen en het begrijpen en produceren van teksten. *Pedagogische Studiën*. 1988, *65*, 32-44.

Bartlett, F.C., *Remembering. A study in experimental and social psychology*. Cambridge: University Press, 1932.

Beaugrande, R. de, *Text production: Toward a science of composition*. Norwood, New Jersey: Ablex Corporation, 1984.

Bereiter, C., Development in writing. In: L.W. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 73-93). Hillsdale, N.J.: Lawrence Erlbaum, 1980.

Bochardt, T., Het schrijfproces: cognitief-psychologisch onderzoek van Flower en Hayes. *Tijdschrift voor Taalbeheersing*, 1984, *6*, 23-42.

Brugman, G.M., Het ontdekken en formuleren van problemen, *Pedagogische Studiën*, 1989, *66*, 23-36.

Claparède, (Ed.), La génèse de l'hypothèse. *Archives de Psychologie*, 1933, *24*, 1-55.

Dewey, J., *How we think*. Boston: Heath, 1910.

Duncker, K., *Zur Psychologie des produktiven Denkens*. (Herdruk: 1963.) Berlin: Springer Verlag, 1935.

Eigler, G., Th. Jechle, G. Merziger & A. Winter, Ueber Beziehungen von Wissen und Textproduzieren. *Unterrichtswissenschaft*, 1987, *15*, 382-395.

Emig, J., *The composing processes of twelfth graders* (Research report no 3). Urbana, Illinois: National Council of teachers in English, 1971.

Faigley. L., R.D. Cherry, D.A. Joliffe & A.M. Skinner, *Assessing writer's knowledge and processes of composing*. Norwood: Ablex Publishing Corporation, 1985.

Flavell, J.H., Metacognition and cognitive monitoring. A new area of cognitive development inquiry. *American Psychologist*, 1979, *34*, 906-911.

Flower, L., *Problem solving strategies for writing*. New York: Harcourt Brace Jovanovich, 1981.

Flower, L. & J.R. Hayes, Problem solving strategies and the writing process. *College English*, 1977, *39*, 449-461.

Flower, L. & J.R. Hayes, The dynamics of composing: Making plans and juggling constraints. In: L.W. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31-50). Hillsdale, N.J.: Lawrence Erlbaum, 1980a.

Flower, L. & J.R. Hayes, The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, 1980b, *31*, 21-32.

Flower, L. & J.R. Hayes, A cognitive process theory of writing. *College Composition and Communication*, 1981a, *32*, 365-387.

Flower, L. & J.R. Hayes, Plans that guide the composing process. In: C.H. Frederiksen & J.F. Dominic, (Eds.). *Writing: The nature, development and teaching of written communication* (pp. 39-58). Hillsdale, N.J.: Lawrence Erlbaum, 1981b.

Flower, L. & J.R. Hayes, The pregnant pause: An inquiry into the nature of planning. *Research in the teaching of English*, 1981c, *15*, 229-243.

Flower, L. & J.R. Hayes, Images, plans and prose. The representation of meaning in writing. *Written communication*, 1984, *1*, 120-160.

Flower, L., J.R. Hayes, L. Carey, K. Schriver & J. Stratman, Detection, diagnosis and the strategies of revision. *College Composition and Communication*, 1986, *37*, 16-55.

Frederiksen, C. H. & J. F. Dominic, Introduction: Perspectives on the activity of writing. In: C. H. Frederikson & J. F. Dominic (Eds.), *Writing: The nature, development and teaching of written communication. Vol. 2: Process, development and communication* (pp. 1-20). Hillsdale, N. J.: Lawrence Erlbaum, 1981.

Frijda, N. H. & J. J. Elshout, Probleemoplossen en denken. In: J. A. Michon, E. G. J. Eykman & L. F. W. de Klerk (Eds.), *Handboek der psychonomie* (pp. 413-446). Deventer: Van Loghum Slaterus, 1978.

Geest, Th. van der, *De computer in het schrijfonderwijs. Perspectief voor een procesbenadering*. Enschede: T. H. Twente, 1986.

Getzels, J. W. & M. Csikszentmihalyi, *The creative vision: A longitudinal study of problem finding in art*. New York: Wiley & Sons, 1976.

Graves, D., Research up date. A new look at writing research. *Language Arts*, 1981, *57*, 913-919.

Greeno, J. G., Natures of problem solving abilities. In: W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol 5) (pp. 239-270). Hillsdale N. J.: Lawrence Erlbaum, 1978.

Groot, A. D. de, *Methodologie. Grondslagen van onderzoek en denken in de gedragswetenschappen*. 's Gravenhage: Mouton & Co, 1961.

Guilford, J. P., *The nature of human intelligence*. New York: McGrawhill, 1967.

Halliday, M. A. K., *Language as social semiotic. The social interpretation of language and meaning*. London: Edward Arnold, 1978.

Hayes, J. R., *Cognitive psychology Thinking and creating*. Homewood: The Dorsey Press, 1978.

Hayes, J. R. & L. Flower, Writing as problem solving. *Visible language*, 1980a, *14*, 388-399.

Hayes, J. R. & L. Flower, Identifying the organization of writing process. In: L. W. Gregg & E. Steinberg (Eds.) *Cognitive processes in writing* (pp. 3-30). Hillsdale, N. J.: Lawrence Erlbaum, 1980b.

Hayes, J. R. & L. Flower, Writing research and the writer. *American Psychologist*, 1986, *41*, 1106-1113.

Hayes, J. R., L. Flower, K. A. Schriver, J. F. Stratman & L. Carey, Cognitive processes in revision. In: S. H. Rosenberg (Eds.), *Advances in applied psycholinguistics*. Vol. 2: *Reading, writing and language learning*. (pp.176-240). Cambridge - New York: Cambridge University Press, 1987.

Hayes-Roth, B. & Fr. Hayes-Roth, A cognitive model of planning. *Cognitive Science*. 1972, *3*, 275-310.

Kintsch, W. & T. A. Van Dijk, Toward a model of text comprehension and production. *Psychological review*, 1978, *85*, 363-394.

Kintsch, W., On modeling comprehension. *Educational Psychologist*, 1979, *14*, 3-14.

Kuhl, J., *Motivation, Konflikt und Handlungskontrolle*. Heidelberg: Springer, 1983.

Lindsay, P. & D. A. Norman, *Human information processing. An introduction to psychology*. San Diego: Academic Press, 1972.

Lowyck, J., E. De Gelder & L. Vanmaele, Schrijven met een computer in het beroepssecundair onderwijs: een pedagogisch-didactische benadering. *VVM*, 1987, *2*, 8-31.

McCutchen, D., Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of memory and language*, 1986, *25*, 431-444.

Miller, G. A., E. Galanter & K. H. Pribram, *Plans and structures of behavior*. New York: Holt & Rinehart, 1960.

Molitor, S., Weiterentwicklung eines Textproduktionsmodell durch Fallstudien. *Unterrichtswissenschaft*, 1987, *15*, 396-409.

Newell, A. & H. A. Simon, *Human problem solving*. Englewood Cliffs, N. J.: Prentice-Hall, 1972.

Newell, A., Reasoning, problem solving and decision processes: The problem space as a fundamental category. In: R. S. Nickerson (Ed.), *Attention and performance* (pp. 693-719). Hillsdale, N. J.: Lawrence Erlbaum, 1980.

Norman, D. A. & D. E. Rumelhart, *Explorations in cognition. Memory and knowledge*. San Francisco: Freeman & Co, 1975.

Rasch, R., Julius Bahle's psychology of musical creation. In: N. H. Frijda & A. D. de Groot (Eds.), *Otto Selz: His contribution to psychology* (pp. 164-191). The Hague - New York: Mouton, s.a.

Rohman, D. G., & A. O. Wlecke, *Prewriting: The construction and application of models for concept formation in writing* (U.S. Office of Education Cooperative Research Project, No. 2174) East Lansing, MT: Michigan State University, 1964.

Rose, M., Rigid rules, inflexible plans and the stiffling of language: A cognitivist analysis of writer's block. College *Composition and Communication*, 1980, *31*, 389-401.

Rose, M., Complexity, rigor, evolving method and the puzzle of writer's block: Thoughts on composing-process research. In: M. Rose (Ed.), *When a writer can't write. Studies in writer's block and other composing-process problems* (pp. 227-260). New York: The Guilford Press, 1985.

Rubin, A. & B. Bruce, *Quill. Reading and writing with a mico-computer* (B.B.N.N. Report No. 5410). Champaign, Ill.: University of Illinois, 1984.

Rumelhart, D. E., Notes on a schema for stories. In: D. Bobrow, A. Collins (Eds.), *Representation and understandig* (pp. 211-236). New York: Academic Press, 1975.

Rumelhart, D. E., P. H. Lindsay, D. A. Norman, A process model for longterm memory. In: E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 198-248). New York: Academic Press, 1972.

Scardamalia, M. & P. Paris, The function of explicit discourse knowledge in the development of text representations and composing strategies. *Cognition and instruction*, 1985, *2*, 1-39.

Scardamalia, M. & C. Bereiter, Research on written composition. In: M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 778-803). New York: Macmillan Publishing Company, 1986.

Scardamalia, M. & C. Bereiter, Knowledge telling and knowledge transforming in written composition. In: Sh. Rosenberg (Eds.), *Advances in applied psycholinguistics*. Vol. 2: *Reading, writing and language learning* (pp. 142-175). Cambridge - New York: Cambridge University Press, 1987.

Schank, R. C., Conceptual dependency: A theory of natural language understanding. *Cognitive psychology*, 1972, *3*, 552 - 631.

Selz, O., *Zur Psychologie des produktiven Denkens und des Irrtums. Eine experimentelle Untersuchung* (4 vols.). Bonn: Cohen, 1922.

Simon, H. A., The structure of ill-structured problems. *Artificial Intelligence*, 1973, *4*, 181-201.

Simon, H. A. & J. R. Hayes, Understanding written problem instructions. In: H. A. Simon (Ed.), *Models of thought* (pp. 451-476). London: Yale University Press, 1974.

Simon, H. A. & J. R. Hayes, The understanding process: Problems isomorphs. *Cognitive Psychology*, 1976, *8*, 165-190.

Sternberg, J., *Beyond I. Q. A triarchic theory of human intelligence*. Cambridge: Cambridge University Press, 1985.

Vanmaele, L., Het lees- en schrijfproces bij zwakkere leerlingen: een literatuurverkenning (Intern rapport 1). Leuven: Departement Pedagogische Wetenschappen K.U. Leuven, 1986.

Vanmaele, L. & J. Lowyck, *Schrijfmodellen en hun achtergronden* (Intern rapport 3). Leuven: Onderzoekscentrum voor Educatieve Technologie. Departement Pedagogische Wetenschappen K.U. Leuven, 1990.

Vanmaele, L., *Schrijfprocessen. Theorie en empirische bevindingen* (Paper gepresenteerd tijdens het SLO-Seminarie) Lattrop, 1988.

Vanmaele, L., *Schrijven bij taalzwakke leerlingen* (Intern rapport 4). Onderzoekscentrum voor Educatieve Technologie, Departement Pedagogische Wetenschappen K.U. Leuven, 1990.

Vos, J., Schrijfprocesonderzoek en schrijfmethodes: het onderzoek van Flower en Hayes en het leerboek van Flower. *Tijdschrift voor Taalbeheersing*, 1983, *5*, (4), 258-269.

Voss, J. F., T. R. Greene, T. A. Post & B. C. Penner, Problem solving skill in the social sciences. In: G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research theory* (Vol. 17) (pp. 165-213). New York: Academic Press, 1983.

Wiener, N., *Cybernetics: on control and communication in the animal and the machine*. New York: John Wiley & Sons, 1948.

Curricula vitae

L. Vanmaele (Onderzoekscentrum voor Educatieve Technologie 'Educatec', K.U. Leuven) is sinds 1986 wetenschappelijk medewerker. Ze richt zich vooral op de studie van schrijfprocessen bij (zwakke) leerlingen en op het ontwerpen van een adequate (computer-)leeromgeving. In dit verband werkt ze deeltijds mee aan het Ministerieel Project 'Functionele Informatica en het aanleren van basisvaardigheden in lezen en schrijven'.

J. Lowyck (gewoon hoogleraar) is het Hoofd van het Onderzoekscentrum voor Educatieve Technologie 'Educatec' en Coördinator van de Universitaire Dienst voor Cursusontwikkeling 'UDC' aan de K.U. Leuven. Speciale aandacht gaat uit naar de constructie van leeromgevingen, met gebruikmaking van onderzoeksuitkomsten uit het domein van de cognitieve psychologie, de linguïstiek en de lerarencognities.

Adres: Educatec, Departement Pedagogische Wetenschappen, Katholieke Universiteit Leuven, Vesaliusstraat 2, B-3000 Leuven

Manuscript aanvaard 2-10-'89

Summary

Vanmaele, L & J. Lowyck. 'The writing model of Hayes and Flower: a critical appraisal.' *Pedagogische Studiën*, 1990, *67*, 205-221.

The writing model of Hayes and Flower is a very dominant one in most research on writing and more recently in the domain of instructional design. In this article, the characteristics of the model are briefly reviewed in order to estimate its suitability as the basis for the construction of a learning environment. First, the aims, methods and results of the model are described. Second, the limitations of the model are revealed, mainly by the use of a contrastive analysis of existing models. Third, the model is evaluated using four criteria: reality representation, generalization, prescriptive value and clarity. In the conclusion, the limitations as well as the strenghts of the model are elucidated in their contribution to the construction of a writing environment for novices.

Differentiatie en selectie in het basisonderwijs*

M. G. WEIDE en G. J. REEZIGT**
RION, Instituut voor Onderwijsonderzoek, Rijksuniversiteit Groningen

Samenvatting

Aan het eind van de basisschool stromen leerlingen door naar verschillende vormen van voortgezet onderwijs. Er is sprake van selectie, wanneer leerlingen in hun doorstroommogelijkheden worden beperkt. In dit artikel wordt ervan uitgegaan dat dit het geval is wanneer leerkrachten voor een leerling niet tenminste communale doelstellingen nastreven.

In het hier gepresenteerde onderzoek staat de relatie tussen differentiatie en (milieu- en seksespecifieke) selectie centraal. Gegevens zijn verzameld bij 158 leerkrachten van groep 8 van de basisschool. De leerkrachten hebben voor zes leerlingen van hun groep en in het algemeen aangegeven in welke mate ze taal- en rekendoelstellingen nastreven. Een doelstelling is communaal genoemd wanneer deze door tenminste 75% van de leerkrachten in hoge mate wordt nagestreefd. Analyses wijzen uit dat er sprake is van milieuspecifieke selectie, maar het differentiatiemodel dat de leerkracht hanteert is hierop niet van invloed. Voor seksespecifieke selectie zijn geen aanwijzingen gevonden.

1 *Inleiding*

Op de bassischool vindt selectie op verschillende manieren en op verschillende momenten plaats. Leerlingen blijven zitten, ze worden verwezen naar het speciaal onderwijs en aan het einde van de basisschool stromen leerlingen door naar verschillende vormen van voortgezet onderwijs. Er zijn aanwijzingen dat de aard van deze beslissingen milieu- en seksespecifiek is (Tesser, 1981; De Jong & Van Veldhuizen, 1984; Meijnen, 1984). Het is mogelijk dat de wijze waarop leerkrachten in hun groep differentiëren hierop van invloed is.

Op het RION is begin 1989 een langlopend onderzoek naar differentiatie op de basisschool afgesloten. Het eerste deel betrof een inventarisatie van de stand van zaken, in het tweede deel zijn de effecten van differentiatie onderzocht. Onderdeel van het tweede deelonderzoek was de vraag naar de relatie tussen differentiatie en selectie. Vragen die in dit artikel aan de orde komen zijn:

1. Bestaat er een relatie tussen het differentiatiemodel dat een leerkracht hanteert en selectie van leerlingen op microniveau?
2. Is er sprake van milieu- en seksespecifieke selectie en zo ja, is deze gerelateerd aan het differentiatiemodel dat de leerkracht hanteert?

2 *Onderzoek naar selectie*

In onderzoek naar selectiemechanismen in het basisonderwijs worden meestal outputvariabelen op mesoniveau als (percentage) zittenblijven, verwijzing naar het speciaal onderwijs en advies voor het vervolgonderwijs gehanteerd. Dergelijk onderzoek geeft indicaties voor het bestaan van milieu- en seksespecifieke selectie.

Leerlingen uit lagere sociale milieus komen met een achterstand het onderwijs binnen en worden vaker naar het speciaal onderwijs verwezen dan leerlingen uit hogere milieus (Meijnen, 1984). Ook stromen leerlingen uit lagere milieus bij gelijke prestaties vaker door naar een lager type vervolgonderwijs in vergelijking met leerlingen uit hogere milieus (Tesser, 1981).

Jongens blijven vaker zitten dan meisjes en worden vaker verwezen naar het speciaal on-

* Dit onderzoek was mogelijk dankzij een subsidie van het Instituut voor Onderzoek van het Onderwijs (SVO), 's-Gravenhage (SVO-project 5020).
** Wij bedanken H. Biemond voor haar medewerking aan dit onderzoek. Tevens bedanken wij P. de Koning voor zijn kritische commentaar bij een eerdere versie van dit artikel.

Pedagogische Studiën 1990 (67) 222-230

derwijs (De Jong & Van Veldhuizen, 1984). Meisjes krijgen daarentegen vaker een te laag advies in verhouding tot hun capaciteiten (Tesser, 1981).

Over de relatie tussen differentiatie en milieu- en seksespecifieke selectie-mechanismen is vrij weinig bekend. Op microniveau is deze relatie nog nauwelijks onderzocht. Twee onderzoeken worden kort genoemd.

Van der Kley (1983) onderzocht het alledaagse schoolse selectieproces, waarbij in drie klassen is gekeken hoe het didactisch handelen van de leerkracht tot selectie leidt. Het didactisch handelen heeft Van der Kley geoperationaliseerd als het geven van beurten. Hij concludeert dat leerkrachten zich aan het begin van het schooljaar al een bepaald beeld van een leerling vormen. Milieu en sekse spelen bij deze beeldvorming een rol. Dit werkt door op het aantal en soort beurten dat een leerling krijgt. Wanneer leerlingen in homogene prestatiegroepen werken, bestaat het gevaar voor het ontstaan van een stromenland in de basisschool: leerlingen uit hoog milieu komen in een academische stroom terecht, leerlingen uit laag milieu in een lagere beroepsstroom.

Jungbluth (1985) wijst op het bestaan van verschillende prestatiestromen, die voorbereiden op respectievelijk het VWO en het LBO. Leerkrachten hebben een bepaald beeld van de prestatiegeschiktheid van leerlingen. Dit beeld is gerelateerd aan het sociaal milieu van de leerling. Het onderwijsaanbod dat een leerling krijgt hangt samen met de inschatting van de leerkracht van de prestatiegeschiktheid. Voor een leerling die geschikt wordt geacht voor het VWO, worden andere leerdoelen nagestreefd dan voor een leerling die geschikt wordt geacht voor het LBO. Leerkrachten verschillen hierin naarmate ze meer leerlinggericht of meer leerstofgericht zijn. Leerlinggerichte leerkrachten streven bij goede leerlingen een hoger prestatieniveau na dan leerstofgerichte leerkrachten, maar bij zwakke leerlingen juist een lager niveau. Ook zetten leerlinggerichte leerkrachten leerlingen uit lager milieu eerder op een lager spoor dan leerstofgerichte leerkrachten. De selectie vindt in dit geval dus plaats op microniveau, al voor het moment waarop het feitelijke advies voor het voortgezet onderwijs wordt gegeven. Wat differentiatie betreft wijst ook Jungbluth op de gevaren van homogene niveaugroepen. Hij vindt geen aanwijzingen voor seksespecifieke selectie.

3 *Relatie differentiatie en selectie*

Tussen differentiatie en selectie bestaat een spanningsverhouding. Selectie houdt altijd differentiatie in, maar differentiatie hoeft niet zonder meer tot selectie te leiden (De Koning, 1983).

De Koning (1983) definieert *selectie* als het onderscheiden van (groepen) leerlingen op grond van bepaalde kenmerken en het op basis van dit onderscheid verwijzen van leerlingen naar verschillende onderwijsprogramma's die hun doorstroommogelijkheden verkleinen c.q. vergroten. Voor leerlingen op de basisschool betreft dit de doorstroom naar verschillende typen vervolgonderwijs. *Differentiatie* wordt hier gedefinieerd als het aanbrengen van verschillen tussen leerlingen ten aanzien van onderwijscomponenten (zoals instructie of leerstofaanbod) op basis van verschillen in prestaties. Deze definitie is gebaseerd op De Koning (1987). Wanneer er bijvoorbeeld verschillende leerstof is voor leerlingen van een verschillend prestatieniveau, is er sprake van differentiatie in leerstofaanbod.

Differentiatie is niet selectief wanneer het gehanteerd wordt als een didactisch middel om bij alle leerlingen dezelfde doelen te bereiken en het dus betrekking heeft op instructie of leertijd (zie b.v. Rupp, 1985). Differentiatie kan wel selectief zijn wanneer het plaatsvindt op de onderwijscomponenten doelstellingen en leerstofaanbod (De Koning, 1983).

De differentiatiemodellen die in Nederland het meeste voorkomen zijn het basisstof-herhalingsstof-verrijkingsstofmodel (BHV-model), het werken in niveaugroepen en individuele differentiatie (Bosveld, Van Dijk, Reezigt & Weide, 1988. Deze modellen zijn ontstaan uit onvrede met het leerstofjaarklassensysteem.

In het leerstofjaarklassensysteem krijgen alle leerlingen in principe dezelfde leerstof. Ze moeten dezelfde opdrachten in dezelfde tijd maken en krijgen allen klassikaal dezelfde instructie. Vaak passen leerkrachten dit strakke model incidenteel wat aan om tempoverschillen tussen leerlingen op te vangen. Snelle leerlingen mogen iets voor zichzelf doen als ze

klaar zijn of krijgen 'meer van hetzelfde'. Zwakke leerlingen krijgen soms wat extra instructie en hulp. Bij het leerstofjaarklassensysteem is dus de instructie en het leerstofaanbod voor alle leerlingen in principe gelijk, maar de doelstellingen die voor een leerling worden nagestreefd kunnen verschillen. Het model is gericht op de gemiddelde leerling, zwakke leerlingen zullen veelal de behandelde stof (nog) niet helemaal beheersen wanneer de hele groep verder gaat met een volgend leerstofonderdeel. De doelstellingen die voor deze leerlingen nagestreefd worden, worden hierbij aangepast. Voor zwakke leerlingen betekent dit dat zij beneden een bepaald basisniveau terecht kunnen komen.

Bij het BHV-model is de leerstof opgedeeld in basisstof, herhalingsstof en verrijkingsstof. Eerst wordt de basisstof klassikaal aangeboden. Leerlingen die de stof hierna nog niet onder de knie hebben, krijgen herhalingsstof; leerlingen die de stof al wel beheersen, krijgen verrijkingsstof. In principe wordt dus voor alle leerlingen gestreefd naar beheersing van de basisstof. Slechts wat de doelstellingen die op de verrijkingsstof betrekking hebben betreft, maken leerkrachten verschil tussen leerlingen. Deze doelstellingen worden voor sommigen wel, voor anderen niet nagestreefd. Prestaties zijn hierbij het criterium. Hoewel het BHV-model de klassikale werkwijze gemeen heeft met het leerstofjaarklassensysteem, wijkt het van dit model essentieel af ten aanzien van de nadruk op het beheersen van basisstof door alle leerlingen en het systematisch inrichten van de lessen met het oog op dit doel.

Bij het werken met niveaugroepen zitten leerlingen in homogene prestatiegroepen die verschillende leerstofinhouden aangeboden krijgen. Voor de verschillende groepjes worden dus ook verschillende doelstellingen nagestreefd.

Bij individuele differentiatie ten slotte, werken alle leerlingen in hun eigen tempo de stof door. Leerlingen zijn dus op hetzelfde moment met verschillende leerstofinhouden bezig en krijgen individueel instructie. Een consequentie is dat sommige leerlingen aan het eind van het schooljaar veel meer stof doorgewerkt hebben en dus ook veel meer doelstellingen bereikt hebben dan andere leerlingen.

Dat sommige leerlingen minder stof krijgen dan andere leerlingen en minder doelstellingen bereiken, hoeft nog niet te betekenen dat er sprake is van selectie. Om hiervan te kunnen spreken is een criterium nodig, waarbij geldt: wanneer een leerling dit leerstofaanbod niet meer krijgt, of wanneer deze doelstelling niet meer voor hem wordt nagestreefd, is sprake van selectie.

De Koning (1983) noemt beperking van de doorstroommogelijkheden van leerlingen selectie. Voor het basisonderwijs is niet zonder meer duidelijk wanneer van een dergelijke beperking sprake is. In het onderzoek dat hier beschreven wordt is om die reden differentiatie *selectief* genoemd wanneer niet voor alle leerlingen tenminste *communale* doelstellingen worden nagestreefd. Communale doelstellingen betreffen de basisstof, stof die alle leerlingen in ieder geval moeten beheersen. De redenering is dan dat beheersing van communale doelstellingen nodig is, wil een leerling na het verlaten van de basisschool toegang kunnen krijgen tot (in principe alle) vormen van voortgezet onderwijs. Wanneer deze doelstellingen voor sommige leerlingen niet worden nagestreefd, worden hun doorstroommogelijkheden beperkt en treedt selectie op.

Voor de basisschool zijn nog geen officiële communale doelstellingen of landelijke eindtermen vastgelegd en de discussie hierover is nog volop gaande. Omdat er bij gebrek aan eindtermen geen duidelijk selectiecriterium voorhanden was, is door de onderzoekers zelf (zie ook par. 4) een indeling gemaakt in communale en differentiële doelstellingen (Weide & Reezigt, 1988).
Verwacht kan worden dat bij het BHV-model de nadruk zal liggen op het bereiken van de communale doelstellingen en dat er wat dit betreft geen of weinig verschil gemaakt wordt tussen leerlingen. Bij niveaugroepen en individuele differentiatie wordt minder nadrukkelijk naar het beheersen van communale doelstellingen gestreefd. Bij deze modellen kan verwacht worden dat er grotere verschillen in nagestreefde doelstellingen zullen ontstaan voor leerlingen van verschillend prestatieniveau.

4 *Instrumenten voor het meten van selectie en differentiatie*

Om te bepalen of leerkrachten selectief te werk gaan, dat wil zeggen communale doelstellingen niet voor alle leerlingen nastreven, moet eerst duidelijk worden welke doelstellingen communaal genoemd kunnen worden. In eerste instantie zijn hiervoor 15 taal- en 15 rekendeskundigen ingeschakeld. Voor de vakken taal en rekenen is een lijst concrete produktdoelstellingen geformuleerd, gebaseerd op materiaal van Bos (1978), Stroomberg (1978), PPON (Wesdorp, Van der Bergh, Bos, Hoeksma, Oostdam, Scheerens & Triesscheijn, 1986) en het CITO (1986). De deskundigen moesten per doelstelling aangeven of zij deze communaal achtten of niet. Het streven was op basis van de overeenstemming van hun oordelen te komen tot een set communale en een set differentiële doelstellingen.

Deze procedure had niet het gewenste resultaat: de deskundigen stemden niet voldoende overeen wat de communaliteit van de doelstellingen betreft (Weide & Reezigt, 1988). Wel zijn suggesties van deskundigen gebruikt voor bijstelling van de lijst. Een aantal (te) specifiek geformuleerde doelstellingen is globaler geformuleerd. Daarnaast zijn doelstellingen samengenomen, zijn voorbeelden bij doelstellingen gewijzigd en is een aantal doelstellingen aan de lijst toegevoegd. De uiteindelijke lijst bestaat uit 35 taal- en 32 rekendoelstellingen. Uit een try-out van de lijst bij 50 leerkrachten kwam naar voren dat de lijst een goede weerspiegeling is van het taal- en rekendomein voor het laatste leerjaar van de basisschool.

De taaldoelstellingen hebben betrekking op verschillende gebieden zoals spreken, (begrijpend) lezen, spellen, stellen, grammatica en het gebruik van informatiemiddelen. Een aantal voorbeelden:

- correct spellen van werkwoordsvormen (b.v. vind jij, hij heeft geboft);
- begrijpen van eenvoudige handleidingen, gebruiksaanwijzingen en mondelinge/schriftelijke instructies (b.v. hoe je een band moet plakken).

De rekendoelstellingen betreffen de onderdelen hele getallen en kommagetallen, breuken en gemengde getallen, procenten, verhoudingen, meten, tijd- en geldrekenen en tabellen en grafieken. Binnen deze onderdelen zijn voor zover mogelijk doelstellingen geformuleerd die betrekking hebben op inzicht, bewerkingen, schatten en vraagstukken. Een aantal doelstellingen heeft betrekking op inzicht in de relatie tussen verschillende rekenonderdelen. Voorbeelden van rekendoelstellingen zijn:

- bewerkingen (optellen, aftrekken, vermenigvuldigen, delen) onder de 100 uit het hoofd uitvoeren;
- inzicht in verhoudingen; vergroten en verkleinen van figuren.

Het instrument is zo opgezet dat leerkrachten van groep 8 voor zes door de onderzoekers naar milieu en sekse geselecteerde leerlingen van hun groep per doelstelling moesten aangeven in welke mate ze deze voor elk van de betreffende leerlingen nastreven. Antwoordmogelijkheden zijn 'niet', 'in geringe mate' en 'in hoge mate.' Daarnaast moesten ze per doelstelling aangeven in welke mate ze deze in het algemeen, voor alle leerlingen, nastreven.

Omdat het vaststellen van communaliteit via deskundigen niet mogelijk was, is besloten de leerkrachten zelf als maatgevend te beschouwen. Een doelstelling is communaal genoemd, wanneer deze door 75% van alle leerkrachten voor alle leerlingen in hoge mate wordt nagestreefd. Wanneer dit niet het geval is, is een doelstelling differentieel genoemd. Op deze wijze is vastgesteld dat voor het vak taal 17 van de doelstellingen communaal en 18 differentieel zijn; voor rekenen zijn er 14 communale en 18 differentiële doelstellingen.
Om vast te stellen op welke wijze leerkrachten differentiëren moesten zij vragen beantwoorden over hun handelwijze ten aanzien van verschillende onderwijscomponenten bij taal en rekenen. Met behulp van de analysetechniek HOMALS voor nominale data is gekeken welke differentiatiemodellen in de antwoordpatronen van leerkrachten teruggevonden worden (Weide, 1988; Reezigt & Weide, 1988, 1989). Naast het leerstofjaarklassensysteem zijn het BHV-model, het werken in niveaugroepen en individuele differentiatie daadwerkelijk te onderscheiden. Niet alle leerkrachten hebben echter een antwoordpatroon dat eenduidig past bij een van deze modellen. Ongeveer een kwart van de leerkrachten hanteert een mengvorm van twee modellen (vooral leerstofjaarklas/BHV). Ruim 10% combineert elementen uit meer dan twee modellen.

5 *Onderzoeksopzet*

Het onderzoek naar de relatie tussen differentiatie en selectie is een onderdeel van een survey-onderzoek naar de effecten van differentiatie op een a-selecte steekproef van 250 basisscholen. Er zijn gegevens verzameld bij twee leerlingcohorten en hun leerkrachten in twee achtereenvolgende leerjaren: in 1987 in groep 5 en 7, in 1988 in groep 6 en 8 (dezelfde leerlingen, andere leerkrachten). Bij de leerlingen zijn IQ-toetsen, prestatietoetsen voor taal en rekenen en attitudevragenlijsten afgenomen.

Het onderhavige onderzoek heeft alleen betrekking op groep 8. De reden hiervoor is dat het niet haalbaar leek voor lagere leerjaren doelstellingensets te construeren waar leerkrachten uit die leerjaren zich zonder meer in zouden kunnen herkennen. Dit was wel haalbaar voor het laatste leerjaar, omdat van een zekere consensus over de einddoelstellingen van het basisonderwijs sprake is.

In 1987 waren bij alle leerlingen gegevens verzameld over hun sociaal milieu en sekse. Het milieu is vastgesteld door de leerkracht te vragen naar het opleidings- en beroepsniveau van beide ouders. In 1988 zijn op basis van deze gegevens uit elke groep 8 door de onderzoekers zes leerlingen geselecteerd, drie jongens en drie meisjes, telkens een uit hoog, een uit gemiddeld en een uit laag milieu.

Vanwege te weinig leerlingen of een ongelijke milieuspreiding waren niet alle groepen geschikt voor een selectie van leerlingen. Uiteindelijk waren 158 leerkrachtlijsten geschikt voor analyse.

6 *Methode*

De gegevens zijn geanalyseerd met het programma VARCL (Longford, 1988), dat rekening houdt met het multi-level karakter van de data. Dit betekent dat er gewerkt wordt met data op twee (of meer) niveaus en dat er niet geaggregeerd of gedisaggregeerd behoeft te worden met alle nadelen die daaraan kleven. Het voordeel van VARCL is dat er op simpele wijze gekeken kan worden naar milieu- en seksespecifieke effecten (zie verder Aitkin & Longford, 1986; Bosker, 1988).

Om het effect van differentiatiemodellen te bepalen is per doelstellingenset gekeken of leerkrachten die een bepaald differentiatiemodel hanteren, verschillen van leerkrachten die binnen het leerstofjaarklassensysteem werken in het percentage doelstellingen dat ze voor hun leerlingen nastreven. De analyses op de communale doelstellingensets geven antwoord op de vraag naar de relatie tussen differentiatie en selectie (vraag 1).

Omdat het, zeker bij de differentiële doelstellingen, voor de hand ligt dat de prestaties van een leerling samenhangen met de nagestreefde doelstellingen, is gekeken of deze relatie inderdaad bestaat. Daarna is naar het effect van de differentiatiemodellen gekeken, gecontroleerd voor prestaties.

Vervolgens is gekeken of de relatie tussen prestaties en doelstellingen van leerkracht tot leerkracht verschilt en zo ja, of het differentiatiemodel een mogelijke verklaring voor dit verschil geeft. Deze zelfde procedure is gevolgd voor de leerlingvariabelen milieu en sekse. Hierbij zijn de prestaties van de leerling telkens als covariabele meegenomen. Deze laatste analyses geven antwoord op de vraag naar milieu- en seksespecifieke selectie (vraag 2).

7 *Resultaten*

7.1 *Aantallen leerkrachten*

Niet alle 158 leerkrachten die een bruikbare doelstellingenlijst hadden ingevuld, konden worden betrokken in de VARCL-analyses. Ten eerste moest een leerkracht in het leerstofjaarklassensysteem of één van de differentiatiemodellen (het BHV-model, werken in niveaugroepen of individuele differentiatie) ingedeeld kunnen worden op grond van zijn antwoordpatroon. Bovendien moest de leerkracht een valide waarde hebben op elke doelstellingenset omdat het VARCL-programma ontbrekende waarden (nog) niet kan hanteren. Deze voorwaarden leidden ertoe dat voor taal en rekenen elk 75 leerkrachten in de analyses zijn opgenomen. De verdeling van deze leerkrachten over de onderwijsmodellen is te vinden in Tabel 1.

De relatieve verdeling wijkt slechts in geringe mate af van de verdeling van leerkrachten in de totale steekproef (Reezigt & Weide, 1989). Opvallend is dat weinig leerkrachten bij taal

Tabel 1 *Aantallen en percentages leerkrachten in de VARCL-analyses, per vak, per model*

	TAAL		REKENEN	
	N	%	N	%
Leerstofjaarklas	50	66,7	15	20,0
BHV-model	18	24,0	34	45,3
Niveaugroepen	3	4,0	12	16,0
Indiv. differentiatie	4	5,3	14	18,7
	75	100	75	100

de modellen niveaugroepen en individuele differentie hanteren. Al eerder bleek (Bosveld e.a., 1988) dat niveaudifferentiatie bij taal relatief weinig voorkomt.

Bij de interpretatie van de uitkomsten moet rekening worden gehouden met het geringe aantal leerkrachten in deze modellen.

7.2 *Relatie differentiatie en selectie*

Om de effecten van differentiatie te onderzoeken en specifiek de relatie tussen differentiatie en selectie, is gekeken of de variantie op leerkrachtniveau, de verschillen tussen leerkrachten die er bestaan in de doelstellingen die ze voor hun leerlingen nastreven, verklaard kan worden door de variabele differentiatiemodel. In feite is in VARCL elk differentiatiemodel afgezet tegen het leerstofjaarklassensysteem en is er dus gekeken of leerkrachten, die volgens een differentiatiemodel werken, significant afwijken van leerkrachten die volgens het leerstofjaarklassensysteem werken. In Tabel 2 wordt een overzicht gegeven.

Tabel 2 *Gemiddeld percentage nagestreefde doelstellingen, per vak, per doelstellingenset, per model*

	TAAL		REKENEN	
	comm.	diff.	comm.	diff.
Leerstofjaarklas	90	69	94	81
BHV-model	91	75	90	78
Niveaugroepen	92	73	93	76
Indiv. differentiatie	96	79	90	74

Uit Tabel 2 komt duidelijk naar voren, en dat ligt ook voor de hand, dat bij beide vakken meer communale doelstellingen worden nagestreefd dan differentiële doelstellingen. De verschillen tussen de vier onderwijsmodellen zijn gering en niet significant.

Het ontbreken van statistisch significante verschillen in gemiddelden op de *communuale* doelstellingensets leidt tot de conclusie dat er geen directe relatie bestaat tussen het differentiatiemodel dat een leerkracht hanteert en de nadruk die hij legt op het nastreven van communale doelstellingen. Differentiatie leidt dus niet tot selectie, zoals dit begrip in dit onderzoek gedefinieerd is.

7.3 *Relatie prestaties en doelstellingen*

Zoals gezegd ligt een relatie tussen het prestatieniveau van een leerling en de doelstellingen die een leerkracht voor deze leerling nastreeft voor de hand. Voor de communale doelstellingen is dat echter in mindere mate het geval dan voor de differentiële doelstellingen, omdat van communale doelstellingen verwacht kan worden dat ze in principe voor alle leerlingen (ongeacht hun prestaties) in dezelfde mate nagestreefd worden. Van differentiële doelstellingen wordt juist verwacht dat ze voor goede leerlingen meer nagestreefd worden dan voor zwakke leerlingen.

De VARCL-analyses laten zien dat prestaties met alle doelstellingensets samenhangen. Dit is wat de communale doelstellingen betreft tegengesteld aan de verwachting. De percentages doelstellingen die leerkrachten voor leerlingen nastreven wanneer gecontroleerd wordt voor prestatieniveau, verschillen enigszins van de percentages in Tabel 2. Een overzicht wordt gegeven in Tabel 3.

Tabel 3 *Gemiddeld percentage nagestreefde doelstellingen, per vak, per doelstellingenset, per model, gecontroleerd voor prestatieniveau*

	TAAL		REKENEN	
	comm.	diff.	comm.	diff.
Leerstofjaarklas	86	67	88	73
BHV-model	86	70	84	74
Niveaugroepen	93	76	90	78
Indiv. differentiatie	96*	76	90	75

* Verschil met leerstofjaarklas significant op 10%-niveau (VARCL-analyses)

De analyses laten zien dat leerkrachten, die bij taal individuele differentiatie toepassen, wanneer gecontroleerd wordt voor prestaties, significant meer communale doelstellingen voor hun leerlingen nastreven dan leerkrachten die volgens het leerstofjaarklassensysteem wer-

ken. Hoe dit geïnterpreteerd moet worden is niet zonder meer duidelijk. Bovendien moet rekening worden gehouden met het feit dat het slechts om een afwijking op 10%-niveau gaat en dat de analyses slechts vier leerkrachten betreffen.
De volgende stap was te kijken of de relatie tussen prestaties en doelstellingen van leerkracht tot leerkracht verschilt en zo ja, of het differentiatiemodel iets van dit verschil verklaart. Uit de VARCL-analyses blijkt dat deze relatie inderdaad voor alle doelstellingensets van leerkracht tot leerkracht verschilt. Slechts in één geval echter kan het differentiatiemodel iets van deze verschillen verklaren. In Tabel 4 worden de correlaties tussen prestaties en doelstellingen gepresenteerd.

Tabel 4 *Correlaties tussen prestaties en doelstellingensets, per model*

	DOELSTELLINGEN			
	TAAL		REKENEN	
	comm.	diff.	comm.	diff.
Leerstofjaarklas	.37	.39	.51	.61
BHV-model	.21	.27	.32	.42*
Niveaugroepen	-	-	.51	.59
Indiv. differentiatie	-	-	.55	.58

* Verschil met leerstofjaarklas significant op 5%-niveau (VARCL-analyses)

De relatie tussen prestaties en doelstellingen is binnen het BHV-model telkens zwakker dan binnen de overige modellen. Alleen bij de differentiële rekendoelstellingen wijkt de relatie significant af van de relatie binnen het leerstofjaarklassensysteem. Anders gezegd, binnen het BHV-model is het prestatieniveau in mindere mate gekoppeld aan de differentiële rekendoelstellingen die leerkrachten nastreven.

7.4 *Milieu, sekse en doelstellingen*

Omdat prestaties aan alle doelstellingensets gerelateerd zijn, is het prestatieniveau van de leerling telkens als covariabele meegenomen in de analyses met milieu en sekse.

Eerst is gekeken of er een relatie is tussen milieu en doelstellingen. Dit is voor alle doelstellingensets het geval. Voor leerlingen uit laag milieu worden minder doelstellingen nagestreefd dan voor leerlingen uit hoog milieu. Wanneer zowel voor prestaties als voor milieu gecontroleerd wordt, is er geen effect van het model.

Voor taal geldt dat de relatie tussen milieu en doelstellingen van leerkracht tot leerkracht verschilt. Differentiatiemodel verklaart echter bij geen van beide doelstellingensets iets van deze verschillen. Bij rekenen is er zelfs geen significant verschil tussen leerkrachten.

De VARCL-analyses wijzen nergens op een relatie tussen sekse en doelstellingen. Theoretisch is het mogelijk dat dit resultaat het gevolg is van het feit dat de ene leerkracht meer doelstellingen voor jongens nastreeft dan voor meisjes, terwijl de andere leerkracht juist meer doelstellingen voor meisjes nastreeft dan voor jongens. Daarom is ook nog gekeken of er verschil is van leerkracht tot leerkracht in de relatie tussen sekse en doelstellingen. Dit bleek niet het geval te zijn. Leerkrachten streven dus voor jongens en meisjes evenveel doelstellingen na.

Samengevat kan gesteld worden dat er weliswaar sprake is van een relatie tussen milieu en nagestreefde doelstellingen, en dus ook van milieuspecifieke selectie, maar niet in samenhang met het differentiatiemodel dat een leerkracht hanteert. Voor seksespecifieke selectie zijn geen aanwijzingen gevonden.

8 *Discussie*

Leerkrachten streven niet voor alle leerlingen in hun groep in dezelfde mate doelstellingen na. Voor goede leerlingen streven ze meer doelstellingen na dan voor zwakke leerlingen. Ook voor leerlingen uit hoog milieu streven ze meer doelstellingen na dan voor leerlingen uit laag milieu, zelfs nadat gecontroleerd wordt voor prestaties.

Voor de communale taal- en rekendoelstellingen is dit resultaat onverwacht. Communale doelstellingen moeten volgens het grootste deel van de leerkrachten in het algemeen, voor alle leerlingen, in hoge mate worden nagestreefd en kunnen worden gezien als de weerspiegeling van een basispakket aan kennis en vaardigheden, dat alle leerlingen bij het verlaten van de basisschool zouden moeten beheersen. Toch maken leerkrachten ook wat deze doelstellingen betreft verschil tussen leerlingen. In feite is hier sprake van selectie, ervan uitgaande dat leerlingen die dit basispakket niet krijgen aangeboden beperkt worden

in hun doorstroommogelijkheden.

Wat de differentiële doelstellingen betreft ligt de zaak anders. Hierbij wordt juist wel verwacht dat er verschil wordt gemaakt tussen leerlingen van verschillend prestatieniveau. Het milieu van leerlingen zou echter geen rol mogen spelen. Wanneer leerkrachten voor leerlingen uit laag milieu minder differentiële doelstellingen nastreven dan voor leerlingen uit hoog milieu, bij een gelijk prestatieniveau, dan is er sprake van milieuspecifiek handelen. Omdat het hier echter niet gaat om communale doelstellingen hoeft dit gegeven nog geen selectie in te houden.

Leerkrachten maken dus verschil tussen leerlingen. Het onderwijsmodel dat ze hanteren is hierop echter nauwelijks van invloed. Er treden slechts incidenteel effecten op, die bovendien niet bij beide onderzochte vakken identiek zijn.

Op grond van de literatuur zou voor niveaugroepen en individuele differentiatie verwacht kunnen worden dat binnen deze modellen de relatie tussen prestaties en – vooral differentiële – doelstellingen sterker zou zijn dan binnen het leerstofjaarklassensysteem, omdat deze modellen juist het meest nadrukkelijk bij prestatieverschillen aansluiten. Dit is echter niet het geval.

Bij deze resultaten moet een aantal kanttekeningen worden geplaatst. Het selectie-instrument meet weliswaar wat leerkrachten zeggen na te streven voor bepaalde leerlingen, maar er is niet gekeken naar de feitelijke leerstof die leerlingen aangeboden krijgen. Dit was gezien de onderzoeksopzet niet haalbaar. Het is natuurlijk mogelijk dat een discrepantie bestaat tussen het nastreven van doelstellingen en het aanbieden van leerstof; in dat geval kan een samenhang tussen differentiatie en selectie op microniveau nog wel degelijk optreden wat het leerstofaanbod betreft, ook al zijn er nu op doelstellingenniveau geen aanwijzingen voor een dergelijke samenhang gevonden. Overigens zijn de doelstellingen zo geformuleerd en met voorbeelden geïllustreerd, dat ze zo dicht mogelijk bij de leerstof aansluiten.

Daarnaast is niet gevraagd aan leerkrachten hoeveel belang zij bij verschillende leerlingen toekennen aan daadwerkelijke beheersing van een doelstelling. Ook op dit punt kunnen leerkrachten, afhankelijk van het onderwijsmodel dat ze hanteren, verschillen. Het daadwerkelijke selectiemoment ligt dan wat later.

Ten slotte moet opgemerkt worden dat het onderzoek zich beperkt tot groep 8, omdat verwacht werd dat er over de einddoelstellingen van het basisonderwijs enige consensus bestaat. Men zou misschien denken dat dan ook de verschillen tussen leerkrachten in de doelstellingen die ze voor hun leerlingen nastreven klein zijn, omdat veel leerkrachten zich juist op deze einddoelstellingen richten en dat dit een reden is voor het feit dat er nauwelijks verschillen tussen de onderwijsmodellen zijn. Dit is echter niet het geval. Er zijn wel degelijk verschillen tussen leerkrachten, alleen zijn kennelijk andere factoren van invloed op de doelstellingen die zij zeggen na te streven.

De conclusie moet vooralsnog luiden dat de onderzochte differentiatiemodellen geen rol van betekenis spelen bij de doelstellingen die leerkrachten voor hun leerlingen zeggen na te streven en dus ook niet meer of minder selectief in de hier omschreven zin zijn dan het leerstofjaarklassensysteem.

Literatuur

Aitkin, M. & N. T. Longford, Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society*, 1986, series *A*, *1-43*, part I.

Bos, D. J., *Empirisch doelstellingenonderzoek van het moedertaalonderwijs*. Amsterdam: RITP, 1978.

Bosker, R. J., *Een hiërarchisch ontwerp voor de evalutie van de basisvorming*. Groningen: RION, 1988.

Bosveld, J. J. F., M. H. van Dijk, G. J. Reezigt & M. G. Weide, Differentiatiemodellen in het kleuter- en lager onderwijs. *Pedagogische Studiën*, 1988, *65*, 250-259.

CITO, *Eindtoets Basisonderwijs Doelenboek*. Inhoudsverantwoording van de Eindtoets Basisonderwijs vanaf 1987. Arnhem: CITO, 1986.

Jong, T. M. de & G. A. van Veldhuizen, *Deelnamecijfers jongens-meisjes speciaal onderwijs*. Haren: RION, 1984.

Jungbluth, P., *Verborgen Differentiatie*. Nijmegen: ITS, 1985.

Kley, P. van der, *Zeg na, jij!* Purmerend: Muusses, 1983.

Koning, P. de, Differentiatie en selectie in het voortgezet basisonderwijs. *Pedagogische Studiën*, 1983, *60*, 122-130.

Koning, P. de, *Programmadifferentiatie in het voortgezet onderwijs*. Lisse: Swets & Zeitlinger, 1987.
Longford, N.T., *Variance component analysis: manual*. Princeton: Educational Testing Service, 1988.
Meijnen, G.W., *Van zes tot twaalf*. Harlingen: SVO, 1984.
Reezigt, G.J. & M.G. Weide, *Differentiatiemodellen op de basisschool: verslaggeving HOMALS-analyses*. Deelrapport I. Groningen: RION, 1988.
Reezigt, G.J. & M.G. Weide, *Effecten van differentiatie*. Resultaten survey-onderzoek. Deelrapport IV. Groningen: RION, 1989.
Stroomberg, H.P., *Communale rekendoelen*. Een empirisch onderzoek naar doelstellingen van het rekenonderwijs. Amsterdam: RITP, 1978.
Tesser, P., *Schoolloopbanenonderzoek in Nederland*. Nijmegen: ITS, 1981.
Weide, M.G., Reconstructing and scaling educational models by means of HOMALS. In: M.G.H. Jansen & W.H. van Schuur (Eds.), *The many faces of multivariate analysis*. Groningen: Rijksuniversiteit, 1988.
Weide, M.G. & G.J. Reezigt, *Differentiatie en selectie: instrumentconstructie*. Deelrapport II. Groningen: RION, 1988.
Wesdorp, H., H. van der Bergh, D.J. Bos, J.B. Hoeksma, R.J. Oostdam, J. Scheerens & B. Triesscheijn, *De haalbaarheid van periodiek peilingsonderzoek: een vooronderzoek op het gebied van het taalonderwijs in de lagere school*. Lisse: Swets & Zeitlinger, SVO-reeks 85, 1986.

Curricula vitae

M.G. Weide (1955) studeerde sociologie (onderwijssociologie en methoden en technieken) aan de Rijksuniversiteit Groningen. Van 1986 tot 1989 verrichtte zij onderzoek naar differentiatie in het basisonderwijs. Sinds 1988 is zij als assistent-in-opleiding werkzaam op onderzoek naar het onderwijsvoorrangsbeleid.

G.J. Reezigt (1955) studeerde onderwijspsychologie aan de Rijksuniversiteit Groningen en verrichtte van 1982 tot 1985 onderzoek naar schoolwerkplanontwikkeling. Vanaf 1984 onderzocht zij differentiatie in het basisonderwijs. Momenteel is zij betrokken bij de evaluatie van het onderwijsvoor- rangsbeleid.

Adres van de auteurs: RION, Instituut voor Onderwijsonderzoek, Rijksuniversiteit Groningen, Postbus 1286, 9701 BG Groningen

Manuscript aanvaard 29-11-'89

Summary

Weide, M.G. & G.J. Reezigt. 'Adaptive instruction and selection in primary education.' *Pedagogische Studiën*, 1990, *67*,222-230.

At the end of primary education pupils go to different types of secondary education. When pupils are restricted in their possibilities to attend these schooltypes, we call this "selection". In this article we state this is the case when teachers do not pursue at least common objectives.

The central theme of this article is the relation between adaptive instruction and selection based on social class and sex. Data were collected from 158 teachers of the last grade of primary school. The teachers indicated for six of their pupils and in general to what degree they pursue language and arithmetic objectives. An objective is called "common" when 75 percent of the teachers highly pursue it. Analyses show that selection based on social class occurs; however, adaptive instruction does not influence this process. No indications were found for selection based on sex.

Vervolgadviezen in relatie tot etniciteit; hoe verhouden prestaties, gezinsachtergrond en leerkrachtoordelen zich tot elkaar?*

P. JUNGBLUTH, CHR. VERHAAK en G. DRIESSEN**
Instituut voor Toegepaste Sociale Wetenschappen (ITS), Nijmegen

Samenvatting

Welke rol spelen leerkrachtoordelen bij de totstandkoming van de vervolgadviezen voor leerlingen uit migrantengroepen? Die vraag wordt beantwoord aan de hand van data die stammen uit effectiviteitsonderzoek naar Onderwijs in Eigen Taal en Cultuur (OETC). Allereerst blijkt dat de 'prestatiesporen' die leerkrachten uitzetten voor hun leerlingen en hun indruk van het 'onderwijsondersteunend thuisklimaat', duidelijk samenhangen met etniciteit. Beide verbanden verdwijnen evenwel volledig wanneer gecorrigeerd wordt voor respectievelijk verschillen in toetsscores en in sociaal-economische statuskenmerken van het gezin. Het verband tussen vervolgadvies en etniciteit kan dus worden herleid tot niet-etnische determinanten; de etnische ongelijkheid in vervolgadviezen verloopt langs mechanismen zoals we die al kennen bij Nederlandse leerlingen. Opnieuw blijkt dat tamelijk arbitraire leerkracht-opvattingen ('welk prestatiespoor voor wie') naast leerprestaties en Sociaal-Economische Status (SES) bepalend zijn voor de vervolgadviezen van alle leerlingen, migranten zowel als andere.

* De analyses die hier worden beschreven, zijn uitgevoerd op verzoek van de Interdepartementale Commissie Minderhedenbeleid en gefinancierd via SVO.

** Onze collega H. Vierke heeft belangrijke bijdragen geleverd voor de gerapporteerde analyses.

1 *Achtergrond*

In het schoolloopbaanonderzoek met betrekking tot migranten in Nederland bestaan tegenstrijdige indicaties over de rol van het leerkrachtadvies bij de overgang van basis- naar voortgezet onderwijs.

Onderzoek naar de relaties tussen milieu en schoolloopbanen heeft meerdere malen het verband gelegd tussen sociaal-economisch milieu en schooladvies (Dronkers, 1987). Bij gelijke prestaties zouden leerlingen uit hogere sociaal-economische milieus hogere schooladviezen krijgen. Dit zou veroorzaakt worden door de leerkracht die zich impliciet laat leiden door sociale stereotypen ten aanzien van economische klasse en onderwijsbestemming.

Ditzelfde mechanisme zou er toe kunnen leiden dat ook allochtone leerlingen bij schooladvisering aan het eind van de basisschool ondergewaardeerd worden: impliciet hanteren leerkrachten bij leerlingen uit culturele minderheidsgroepen maatstaven die discriminerend werken voor het schooladvies dat de betrokken leerling op basis van zijn of haar prestaties krijgt. In onderzoek van Van Esch (1982) wordt dit bevestigd: Turkse en Marokkaanse leerlingen krijgen twee maal zo vaak een LBO-advies dan Nederlandse kinderen met vergelijkbare schoolprestaties.

Hetzelfde mechanisme zou ook nadelig kunnen werken voor de schooladviezen van meisjes. Blok en Saris (1979) laten echter zien dat meisjes bij een vergelijkbare schoolprestatie iets vaker een hoog advies krijgen dan jongens. In plaats van negatieve lijkt hier sprake van positieve discriminatie. Dit beeld wordt bevestigd door De Jong en Van Batenburg (1984) en De Jong (1987). Meisjes krijgen bij een vergelijkbaar IQ een hoger advies dan jongens. In dezelfde publikatie laten De Jong en Van Batenburg zien dat deze praktijk van 'over-advisering' ook van toepassing is op allochtone leerlingen. Onder constanthouding van IQ krijgen leerlingen uit culturele minder-

heidsgroepen een hoger advies dan Nederlandse leerlingen. Dit verband constateren zij het sterkst bij verbale IQ, in iets mindere mate bij de score op de nonverbale test.

Tesser, Van der Werf, Mulder en Weide (1990) stellen in de voorstudie van de evaluatie van het onderwijsvoorrangsbeleid vooralsnog nauwelijks verschil vast tussen de schooladviezen voor allochtone leerlingen en Nederlandse arbeiderskinderen. Binnen beide groepen is sprake van meer en minder kansrijke categorieën waartussen de scheidslijnen langs identieke sociaal-economische gezinskenmerken lopen. Pas in het hoofdonderzoek kan worden vastgesteld of deze resultaten gecorrigeerd moeten worden voor werkelijke schoolprestaties; als dat het geval is, zijn er factoren in het spel die de vervolgadvisering bij minderheden doen afwijken.

Everts, Golhof, Stassen en Teunissen (1986) dragen een alternatieve verklaring aan voor veronderstelde 'over-advisering van allochtone leerlingen'. Zij zoeken die niet bij een extra positieve beoordeling van leerkrachten maar bij de culturele bias van de intelligentie-testen: de leerkrachten zijn in die opvatting minder biased dan de toetsen. Luykx en Uniken Venema (1985) vestigen in eenzelfde verband de aandacht op het hogere aspiratieniveau van met name Turkse leerlingen. Dit zou verklaard kunnen worden door de migratie-achtergrond van deze kinderen.

Al met al voldoende reden om de achtergronden van de adviespraktijk bij migranten kritisch onder de loep te nemen.

2 *Probleemstelling*

In welke mate spelen binnenschoolse factoren een rol bij de totstandkoming van de (overwegend lage) vervolgadviezen die leerlingen uit migrantengroepen krijgen aan het einde van de basisschool? Dat is de overkoepelende vraag voor een reeks van analyses, waarvan de hier gerapporteerde een deel vormen. In die analyses is de vraag toegespitst op de rol van leerkrachtpercepties: welke rol speelt de indruk die leerkrachten hebben van leerlingprestaties en gezinsachtergrond (naast feitelijke prestaties en feitelijke achtergrond) bij de hoogte van het vervolgadvies dat die leerlingen uiteindelijk krijgen?

Deze meer toegespitste vraagstelling bouwt voort op een reeks van ITS-onderzoeken waarin de rol van leerkrachtpercepties centraal staat (vgl. Jungbluth, 1985; Van der Hoeven-Van Doornum & Jungbluth, 1988). Het gaat vooral om het beeld dat de leerkracht heeft van een leerling en het (in reactie daarop en vermoedelijk in interactie met bepaalde beroepsopvattingen) door de leerkracht informeel voor bepaalde typen leerlingen adequaat geachte prestatiespoor. Dit wordt per individuele leerling vastgesteld en indiceert het niveau van leerdoelen dat de leerkracht adequaat vindt.

De vraagstelling past ook in het streven om sociale ongelijkheidsmechanismen in het onderwijs op te sporen die op meerdere typen van sociale ongelijkheid (sekse, milieu, etniciteit) van toepassing zijn (Jungbluth, 1987).

3 *Databestand*

De gegevens die voor de hier beschreven analyses worden gebruikt, zijn oorspronkelijk verzameld ter wille van een effectiviteitsonderzoek naar het Onderwijs in Eigen Taal en Cultuur (Driessen, De Bot & Jungbluth, 1989). Zij zijn afkomstig van ruim 2300 leerlingen en hun leerkrachten uit groep acht, die op honderdtwintig basisscholen zitten waar in principe Onderwijs in Eigen Taal en Cultuur wordt gefaciliteerd. Dat laatste maakt dat de scholen niet representatief zijn voor het Nederlandse basisonderwijs, maar waarschijnlijk juist wel voor dat deel van het basisonderwijs waarin leerlingen uit migrantengroepen zitten die in aanmerking komen voor OETC.

Voor de onderzoekspopulatie heeft dat enkele belangrijke gevolgen:

- de Nederlandse leerlingen zijn waarschijnlijk niet representatief: er is sprake van een oververtegenwoordiging van kansarme groepen (arbeiderskinderen) zoals die typisch is voor scholen met relatief veel migrantenleerlingen (vgl. Jungbluth, Verhaak & Vrieze, 1989);
- Surinaamse en Antilliaanse leerlingen blijken te zijn ondervertegenwoordigd: kennelijk moet er verschil worden gemaakt tussen scholen met relatief hoge concentraties OETC-gerechtigde migrantengroepen en scholen met relatief hoge concentraties van de overige migrantengroepen.

Het totale databestand waarover de hier

gepresenteerde analyses zijn uitgevoerd bestaat uit 2363 leerlingen, waarvan er bijna 1600 een Nederlandse achtergrond hebben. De laatste zijn de klasgenoten van de voor de oorspronkelijke probleemstelling relevante migranten.

Uitgebreide gegevens over de omvang en opzet van het bestand, over de betrokken onderzoeksopzet en over de manier waarop een aantal variabelen zijn geconstrueerd staan vermeld in het onderzoeksverslag over de effectiviteit van het OETC (Driessen, De Bot & Jungbluth, 1989).

4 *Variabelen en analyse-opzet*

In de hier gepresenteerde analyses worden de volgende variabelen gebruikt:

afhankelijk (te verklaren):

- het vervolgadvies dat de leerling uiteindelijk ontving;

onafhankelijk (eerste verklarende factor):

- de etnische achtergrond;

covariaten ('constant' te houden):

- het herkomstmilieu (SES; geïndiceerd door opleiding en beroep van de ouders);
- de prestaties voor Nederlandse taal (CITO-gelegenheidstoets);
- de prestaties voor rekenen (CITO-gelegenheidstoets);

verdere onafhankelijke (concurrerende verklarende factoren):

- het 'prestatiespoor' dat de leerkracht adequaat acht voor de individuele leerling (vijfpuntsschaal; 'ik bied alleen stof aan die dit kind later nodig heeft', 3 items, alfa: 0.73);
- het beeld dat de leerkracht heeft van het 'onderwijsondersteunend thuisklimaat' (vijfpuntsschaal; 'thuis krijgt deze leerling veel ondersteuning en stimulansen', 3 items, alfa: 0.83).

De analyse-opzet is als volgt:

- eerst wordt nagegaan in welke mate de beide 'binnenschoolse' factoren (leerkrachtbeelden) overlappen met de etnische identiteit ('Discrimineert' de leerkracht naar etniciteit?);
- vervolgens wordt – stapsgewijs – nagegaan in welke mate het verband tussen etniciteit en vervolgadvies herleid kan worden tot prestatieverschillen en SES en in welke mate de onderzochte binnenschoolse factoren (leerkrachtoordelen) uiteindelijk het verband tussen advies en etniciteit verklaren.

In deze opzet zijn enkele belangrijke beperkende veronderstellingen vervat:

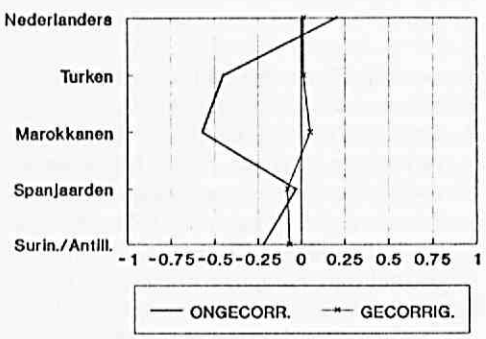
1. Door SES als 'buitenschoolse factor' (covariaat) op te nemen, wordt het 'gewone' verband tussen advies en milieu hier opgevat als een niet-beïnvloedbaar gegeven. Dat is in werkelijkheid niet zo, maar de onderzoeksopzet laat enerzijds weinig keuze, anderzijds wordt de vraag naar de rol van etniciteit door deze werkwijze extra aangescherpt.
2. De via toetsscores geïndiceerde leerlingprestaties worden eveneens als 'buitenschoolse determinanten' (covariaten) opgevat. In werkelijkheid gaat het om een mengeling van talent en o.a. eerdere schoolse invloeden, waarbij de laatste mogelijk verschillen naar etniciteit. Opnieuw een aanscherping van de vraag in welke mate etniciteit een eigen rol speelt in de totstandkoming van het VO-advies (en niet in de voorafgaande onderwijsprocessen).

5 *Beelden die leerkrachten hebben van migranten*

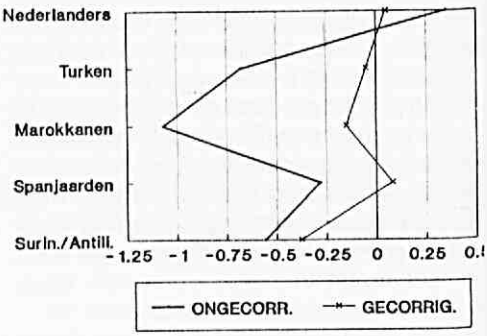
'Hoezeer verschillen de prestatiesporen die leerkrachten van toepassing achten voor leerlingen uit verschillende etnische groepen?' Dat is de eerste vraag die hier wordt beantwoord. Daartoe wordt eerst etniciteit in verband gebracht met dat prestatiespoor, waarna dat verband vervolgens wordt gecontroleerd voor de toetsscores van de leerlingen. Die controle houdt in, dat de vraag wordt opgeworpen, in welke mate de verschillen die leerkrachten hier maken, herleid kunnen worden tot verschillen in getoonde leerprestaties.

In Figuur 1 worden de resultaten van de betrokken analyse gepresenteerd. De aanvankelijk aanmerkelijke verschillen (tot meer dan een halve standaarddeviatie) in prestatiesporen tussen onderscheiden etnische groepen, verdwijnen na controle voor toetsscores geheel. De percentages verklaarde variantie zijn als volgt:

- oorspronkelijk (ongecorrigeerd) verband: 7,2% variantie;



Figuur 1 *'Prestatiespoor' naar etnische groep; afwijkingen van het gemiddelde (gem. = 3,1; s.d. = 1,2; % var.: 7,2, resp. 0)*



Figuur 2 *'Onderwijsondersteunend thuisklimaat' volgens leerkracht naar etnische groep; afwijkingen van het gemiddelde (gem. = 3,0; s.d. = 1,2; % var.: 19, resp. 0)*

- na controle voor toetsscores (die 36% variantie verklaren): 0%.

Een vergelijkbare opzet wordt gevolgd ter beantwoording van de vraag: 'Hoezeer varieert de indruk die leerkrachten hebben van het 'onderwijsondersteunend thuisklimaat' met de etniciteit van dat gezin?' In de betrokken analyses wordt in tweede instantie de Sociaal-Economische Status van het gezin constant gehouden, zodat de vraag wordt vertaald naar: 'Is de indruk die leerkrachten hebben van het 'onderwijsondersteunend thuisklimaat' behalve van de 'hoogte van het milieu' ook afhankelijk van de etniciteit?'

In Figuur 2 worden de resultaten van deze analyses weergegeven. Opnieuw verdwijnen de oorspronkelijk forse (tot meer dan 1 s.d.) verschillen (vrijwel) volledig; de op het oog resterende negatieve indruk van Surinaams-Antilliaanse gezinnen is niet significant. De percentages verklaarde variantie zijn als volgt:

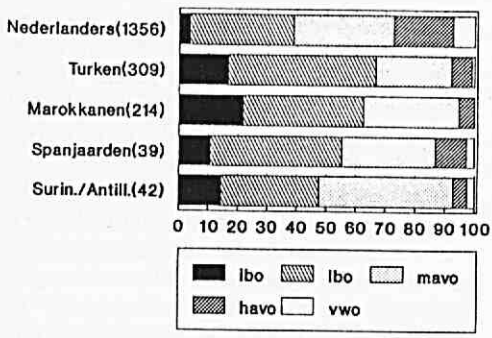
- oorspronkelijk (ongecorrigeerd) verband: 19,4% variantie;
- na controle voor SES (die 35,7% verklaart): 0,3% bij p = .032.

Concluderend kan worden vastgesteld dat leerkrachten alleen in eerste instantie aanmerkelijk verschillen in hun beeld van leerlingen uit onderscheiden etnische groepen en vooral in hun beeld van de herkomstgezinnen van die leerlingen. In het eerste geval ('prestatiesporen') gaat het om verschillen die voliedig herleid kunnen worden tot verschillen in prestatiescores en die dus niet typisch zijn voor etniciteit. Leerlingen uit onderscheiden etnische groepen worden dus op onderscheiden prestatiesporen gezet, zij het dat die verschillen herleid kunnen worden tot eigenschappen waarop ook Nederlandse leerlingen in dat verband 'gesorteerd' worden: het is geen verschijnsel dat typisch is voor etniciteit.

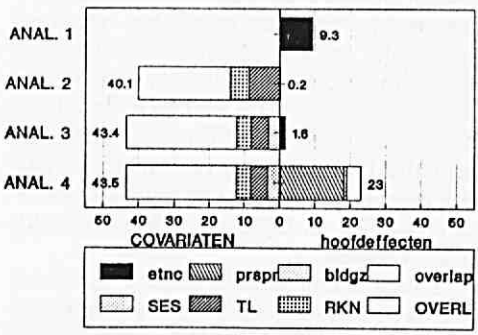
Iets soortgelijks geldt voor het beeld dat de leerkracht heeft van het 'onderwijsondersteunend thuisklimaat' van leerlingen uit onderscheiden etnische groepen. Dat beeld is voor sommige groepen opmerkelijk negatief (vergelijkbaar met een correlatie van ruim .40), maar opnieuw kan dit worden herleid tot mechanismen die ook gelden voor Nederlandse leerlingen: het is de sociaal-economische status van het herkomstgezin die de etnische verschillen wegverklaart. Het negatieve beeld dat leerkrachten in dit verband hebben, is te herleiden tot niet-etnische kenmerken die ook bij Nederlandse leerlingen vergelijkbaar werken (vergelijkbaar met een correlatie van .60).

6 *Determinanten van het vervolgadvies*

In Figuur 3 wordt voor de onderhavige populatie in beeld gebracht hoe de vervolgadviezen verschillen naar etnische groep. In een variantie-analytische opzet (Figuur 4, analyse 1 en Figuur 5 lijn 1) blijkt dat 9,3%



Figuur 3 *Vervolg-advies naar etnische groep; percentuele verdeling per herkomstland.*



Figuur 4 *Determinanten van het V.O.-advies bij migranten; percentages verklaarde variantie in vier opeenvolgende analyses.*

van de verschillen in V.O.-advies voorspeld kan worden op grond van etniciteit (verschillen tot driekwart s.d.).

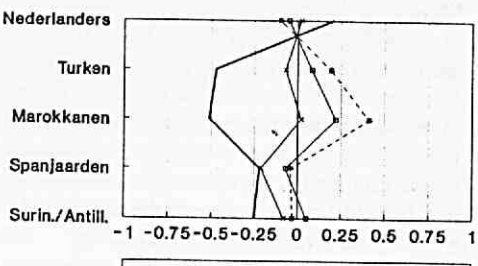
In een trits van steeds complexere analyses (Figuur 4, analyse 2-4 met de bijbehorende correctielijnen 1-3 in Figuur 5) wordt nu bezien of de oorzaak van lage adviezen voor leerlingen uit migrantengroepen moet worden gezocht in andere dan etnische factoren.

In eerste instantie wordt het verband tussen advies en etniciteit (etnc) gecontroleerd voor de toetsscores (taal (TL) en rekenen (RKN)) van de betrokken leerlingen (Figuur 4, analyse 2 en Figuur 5, correctie 1). Deze toetsscores verklaren ruim 40% van het advies, waarbij etniciteit zijn relatie met advies verliest: de verschillen naar etnische groep vallen dus samen met verschillen in schoolprestaties.

Vervolgens, (Figuur 4, analyse 3 en Figuur 5, correctie 2) wordt het verband tussen advies en etniciteit behalve voor toetsscores ook gecontroleerd voor de sociaal-economische status van het gezin (SES). Dit verhoogt uiteraard de verklaarde variantie: het reeds constant gehouden variantie-deel in de adviezen stijgt naar ruim 43%. Dat dit percentage stijgt is een uitvloeisel van het feit dat de leerkracht, bovenop verschillen in getoonde prestaties in de advisering een bekende milieu-'correctie' toevoegt: in het nadeel van de lagere, in het voordeel van de hogere milieus.

Opmerkelijk is dat na deze correctie de etniciteit weer enigszins van belang wordt (ca. 1,5%). De oorzaak (vgl. correctie 2 in Figuur 5) voor deze, weliswaar minieme stijging lijkt te liggen bij de Marokkaanse groep, die – gegeven haar sociaal-economische status – onverwacht hoge(!) adviezen krijgt.

In derde en laatste instantie wordt nu het antwoord gezocht op de hoofdvraag: 'Welke rol is weggelegd voor de onderwijsinterne factoren, het 'prestatiespoor' (prspr) waarop de leerkracht de leerlingen zet en de mate waarin de leerkracht een 'onderwijsondersteunend thuisklimaat' (bldgz) aanwezig acht?' In deze analyse (Figuur 4, analyse 4 en Figuur 5, correctie 3) wordt opnieuw eerst ruim 43% van de variatie in advies 'wegverklaard' door de buitenschoolse factoren 'toetsscores' en SES. Vervolgens wordt nu additioneel nog bijna



Figuur 5 *Het V.O.-advies naar etnische groep, ongecorrigeerd en gecorrigeerd voor prestaties (1), prestaties en SES (2) en prestaties, SES en leerkrachtpercepties (3) (gem. = 2,7; s.d. = 1; % var.: resp. 9, 0, 2 en 0)*

23% variantie verklaard door de binnenschoolse factoren, waarbij de invloed van etniciteit opnieuw nul nadert. Binnen de binnenschoolse factoren is het het 'prestatiespoor' dat vrijwel alle nieuwe verklaringskracht (uniek: 17,9%, incl. overlap: 21,8%) naar zich toe trekt. De leerkracht houdt aan het eigen idee van een bepaalde geschiktheid van de leerling vast bij het uiteindelijk gegeven advies, los van feitelijke prestatieverschillen en feitelijke of veronderstelde gezinskenmerken.

Zo resteert uiteindelijk niets ter verklaring van de etnische ongelijkheid in vervolgadvisering, dat niet ook geldt voor de Nederlandse leerlingen. De verschillen in adviezen kunnen worden herleid tot feitelijke prestatieverschillen en verschillen in SES en tot leerkrachtbeelden die daarop een reactie vormen. Een aanvullende verklaring via etnische factoren lijkt daarna overbodig.

7 *Discussie*

De achterstanden van migrantenleerlingen in het Nederlandse onderwijs komen duidelijk tot uitdrukking op het moment van de overstap van basis- naar voortgezet onderwijs. Het was onze vraag of er daarbij sprake is van binnenschoolse mechanismen die een eigen verklaring vormen voor deze etnische ongelijkheid.

De gepresenteerde analyses maken duidelijk dat er nauwelijks of geen 'eigen' verklaring nodig is voor het verband tussen advieshoogte en etniciteit: de etnische ongelijkheid in VO-advisering is volledig herleidbaar tot die mechanismen die ook bij de Nederlandse leerlingen de VO-advisering (en de daarin tot uitdrukking komende kansenongelijkheid) bepalen. De etnische verschillen in vervolgadviezen kunnen in eerste instantie volledig worden herleid tot prestatieverschillen. Houden we – behalve met prestatieverschillen – ook nog rekening met de standaard-verdiscontering en herkomstmilieu door leerkrachten (lees: lager milieu lager, hoger milieu hoger adviseren dan op grond van prestatieverschillen logisch is), dan lijkt er zelfs sprake van een zekere bevoordeling: Marokkaanse leerlingen krijgen iets hogere adviezen, dan ze gezien hun SES mochten verwachten.

Opmerkelijk zijn de bevindingen over de rol van het 'prestatiespoor' dat de leerkrachten voor de betrokken leerlingen adequaat vinden, bij het uiteindelijke advies. Dit is een bevestiging van eerdere bevindingen (Jungbluth, 1985; Van der Hoeven & Jungbluth, 1989): een eenmaal adequaat geacht 'prestatiespoor' (dat in de voorafgaande onderwijsperiode zelf al van invloed is op de werkelijke prestaties) wordt ook daarna nog consequent doorgetrokken naar het vervolgadvies. Of deze adviespraktijk uitdrukking is van werkelijke professionaliteit (de leerkracht ziet meer dan de toets meet) of van het tegendeel daarvan (toevallige 'pedagogieën' en plaatselijke of generatiegebonden subculturen), dat is vooralsnog onduidelijk, maar dat hier sprake is van een effectiviteitsdeterminant, is wel duidelijk.

Voor het beleid inzake migranten vormen de gerapporteerde onderzoeksbevindingen een zoveelste aanwijzing, dat kansenverbetering in het onderwijs langs de weg moet gaan die ook geldt voor Nederlandse kansarme groepen. Prestatieverbetering (sic) is daarbij prioriteit één en die kan wellicht mede worden bereikt door leerkrachten andere 'prestatiesporen' te doen uitzetten voor achterstandsgroepen. Zowel die 'prestatiesporen' als de standaard 'milieu-correctie' die leerkrachten toepassen bij hun advisering, spelen voor de onderwijskansen van achterstandsgroepen een grote rol. Zo groot, dat het de vraag is of die oncontroleerbare leerkrachtmarges niet teruggedrongen moeten worden ten gunste van empirisch onderbouwde standaarden.

Literatuur

Blok, H. & W. E. Saris, Relevante variabelen bij het doorverwijzen na de lagere school, een structureel model. *Tijdschrift voor Onderwijsresearch*, 1980, *5*, 63-79.

Driessen, G., K. de Bot & P. Jungbluth, *De effectiviteit van het Onderwijs in Eigen Taal en Cultuur. Prestaties van Marokkaanse, Turkse en Spaanse leerlingen*. Nijmegen: ITS, 1989.

Dronkers, J. Onderwijs en sociale ongelijkheid. In: J. van Kemenade, N. Lagerweij, J. Leune & J. Ritzen (red.), *Onderwijs, Bestel en beleid*. Groningen: Wolters-Noordhoff, 1987.

Esch, W. van, *Etnische groepen en het onderwijs*. Den Haag: SVO, 1982.

Esch, W. van, *Toetsprestaties en doorstroomadviezen van allochtone leerlingen in de zesde klas van de lagere scholen*. Nijmegen: ITS, 1983.
Everts, E., A. Goldhof, P. Stassen & J. Teunissen, *De kultureel-etnische situatie op OVB-scholen*. Utrecht: RUU, 1986.
Hoeven-Van Doornum, A. & P. Jungbluth, Leerkrachtverwachtingen, streefniveaus en het advies voor voortgezet onderwijs. *School en Begeleiding*, 1988, *5*, 20, 11-13.
Jong, M.-J. de & Th. van Batenburg, Etnische herkomst, intelligentie en schoolkeuzeadvies. *Pedagogische Studiën*, 1984, *61*, 362-371.
Jong, M.-J. de & J.J.A. Tacq, Het onderwijsniveau van allochtonen leerlingen. *Tijdschrift voor Onderwijsresearch*, 1985, *10*, 130-140.
Jong, M.-J. de, *Herkomst, kennis en kansen. Allochtone en autochtone leerlingen tijdens de overgang van basis- naar voortgezet onderwijs*. Lisse: Swets en Zeitlinger, 1987.
Jungbluth, P., De rivaliteit tussen brede onderwijsdoelen en optimale onderwijskansen. In: Th. Bergen, J. Giesbers & C. Morsch (red.), *Professionalisering van onderwijsgevenden*. Lisse: Swets en Zeitlinger, 1987.
Jungbluth, P., *Verborgen Differentiatie. Leerlingbeeld en onderwijsaanbod op de basisschool*. Nijmegen: ITS, 1985.
Jungbluth, P., Chr. Verhaak & G. Vrieze, *Beleidsadvies formatie-regeling en onderwijsvoorrang*. Nijmegen: ITS, 1989.
Luykx, M. & P. Ulriken Venema, *Wegblijven van school*. Utrecht: RUU, 1985.
Tesser, P., G. van der Werf, L. Mulder & M. Weide, *De beginfase van het Voorrangsbeleid in het Basisonderwijs*. Nijmegen/Groningen: ITS/RION, 1990.

Curricula vitae

P. Jungbluth (1949), *Chr. Verhaak* (1959) en *G. Driessen* (1953) zijn als wetenschappelijk onderzoekers werkzaam in de themagroep 'Onderwijskansen' van het Instituut voor Toegepaste Sociale Wetenschappen (ITS) in Nijmegen.

Adres: ITS, Instituut voor Toegepaste Sociale Wetenschappen, Graafseweg 274, 6532 ZV Nijmegen.

Manuscript aanvaard 16-11-'89

Summary

Jungbluth, P., Chr. Verhaak & G. Driessen. 'Teachers' recommendations regarding secondary education for pupils from ethnic minorities: the influence of achievement, family-background and teachers' assessments.' Pedagogische Studiën. 1990, *67*, 231-237.

What role do teachers' assessments play in the construction of recommendations regarding future secondary education for their pupils from ethnic minorities? At first, the 'achievement tracks' implicitly constructed by teachers and their impression of the 'education-supportive-home-climate' seem evidently related to ethnic background. Both relations however, disappear completely when the data are corrected for the respective differences in pupils' test-scores and social economic status-characteristics of the family. The relation between the recommendations regarding future secondary education and ethnic background can thus be reduced to non-ethnic determinants; the ethnic inequality in these recommendations proceeds through the same mechanisms as the ones already familiar to us from Dutch pupils.

Once again, the fairly arbitrary assessment by teachers ('which achievement-track for which pupil?') appears to determine, together with the pupils' achievements and SES, the recommendations regarding future secondary education for all pupils, including pupils from ethnic minorities.

Boekbespreking

C. Teyken, *Ontwerpen als proces*. Een exploratief onderzoek naar ontwerpprocessen van beginners vanuit onderwijskundig perspectief. Swets & Zeitlinger, Lisse, 1988, 198 pp., *f*35, –, ISBN 90 265 0878 6.

In augustus 1986 overleed na een kortstondige ziekte de onderwijskundige Cees Teyken. Hij was als docent verbonden aan de nieuwe lerarenopleiding Ubbo Emmius te Leeuwarden en bereidde naast dit werk een dissertatie voor over het leren ontwerpen van produkten van grafische en industriële vormgeving. Hij kon zijn nagenoeg gereedgekomen manuscript niet meer aanbieden voor de promotie. Dank zij de inspanningen van zijn echtgenote, vrienden en promotoren is het voorgenomen proefschrift alsnog gepubliceerd in december 1988. Daarmee heeft toch nog een afronding plaatsgevonden van het werk van Teyken. Onderwijspraktijk en -wetenschap kunnen blij zijn met dit initiatief, omdat het werk een groot aantal creatieve oplossingen bevat voor het onderzoeken van het zo moeilijke terrein van de creatieve probleemoplossingsprocessen en tegelijkertijd de onderwijspraktijk concrete handreikingen biedt voor kwalitatief goed ontwerponderwijs.

Deze recensie is een eerbewijs aan Teyken hetgeen tot uitdrukking komt in een samenvatting van de hoofdstukken van het boek met daaraan gekoppeld een kritische bespreking.

In het korte, inleidende *eerste hoofdstuk* wordt het ontwerpprobleem geïntroduceerd en gezien als een slecht gedefinieerd probleem. Het ontwerpen wordt omschreven als 'het ontwikkelen van een nieuw(e) produkt, systeem of situatie opdat bepaalde functies zo optimaal mogelijk worden vervuld' (blz. 10). Afhankelijk van de aard van de oplossing onderscheidt Teyken verschillende vormen van ontwerpen: produktontwerpen, systeemontwerpen en omgevingsontwerpen. Een gehanteerd beginsel is dat ontwerpen principieel verschilt van onderzoeken. Twee beperkingen brengt hij aan: a) Het onderzoek beperkt zich tot een functionele benadering van het produktontwerpen met produkten van grafische en industriële vormgeving; b) de invalshoek voor het ontwerpen is die van het cognitieve proces.

Naar mijn smaak verontachtzaamt de gekozen definitie van 'ontwerpen' de relaties die bestaan tussen een produkt en een systeem of situatie waarin een dergelijk produkt veelal wordt gebruikt. Ten aanzien van een te ontwerpen object zijn niveaus van complexiteit te onderscheiden (De Boer, 1989). Eén van de eerste activiteiten van de ontwerper is de plaats van 'zijn' object in de hiërarchie te bepalen en daarbij direct te zien met welke naastliggende niveaus hij rekening dient te houden. Voorts lijkt me het benadrukken van de verschillen tussen ontwerpen en onderzoeken vooral samen te hangen met het gekozen vakgebied. In andere ontwerpdomeinen – b.v. onderwijskundig ontwerpen; technisch ontwerpen – is het juist van belang waar de twee op zich verschillende handelingen op elkaar aansluiten. Overigens, Teykens boek zelf is een mooi voorbeeld hoe met behulp van een ontwerpprodukt onderzoek wordt bedreven en omgekeerd.

Helaas blijft het geheel van vraagstellingen van de studie vaag omschreven in het eerste hoofdstuk. Het zal vooral gaan om cognitieve processen van ontwerpprocessen. Uit de ondertitel wordt duidelijk dat het gaat om een exploratief onderzoek naar ontwerpprocessen van beginners vanuit onderwijskundig perspectief. Beter ware het geweest het geheel van hoofd- en subvraagstellingen – in hoofdstuk 5 over de evaluatieopzet worden specifieke vraagstellingen geformuleerd – in dit hoofdstuk systematisch op een rij te zetten. Daarmee was direct een goede oriëntering op de hoofdstukken van het boek gegeven. Ik vermoed dat Teyken juist aan deze afronding niet is toegekomen.

In het *tweede hoofdstuk* wordt een literatuursamenvatting gegeven van de resultaten van descriptief onderzoek naar ontwerpprocessen (veelal bij architectuur-experts) in de jaren 1970 - 1980. Teyken kiest vanuit een procesoriëntatie ten aanzien van het onderwijs voor een procesmodel voor ontwerpprocessen. De iteratieve samenhang van het analyseren, synthetiseren en avalueren ziet hij als fundamentele operaties van ontwerpprocessen. De volgorde van uitvoering en de inhoud waarop deze fundamentele operaties zijn gericht, variëren met de aard van de ontwerpproblemen en het stadium waarin het ontwerpproces zich bevindt. 'Analyseren' is met name gericht op het bepalen van de functies die het eindprodukt dient te bezitten en op het onderkennen van de deelproblemen. Het begrip 'synthese' verwijst naar het produceren en combineren van ideeën en oplossingen en het begrip 'evaluatie' betreft het bepalen van de waarde van ideeën en oplossingen. Op grond van een bespreking van enkele procesmodellen voor ontwerpprocessen komt de auteur tot een eigen fasenmodel met tussenprodukten waarin rekening is gehouden met gebruik in de onderwijspraktijk en met diverse cognitieve operaties die veelal in de probleemoplossingsliteratuur worden genoemd. Dit fasenmodel zal

worden gebruikt als interpretatiemodel voor de kwantitatieve en kwalitatieve protocolanalyse.

Ik heb de indruk dat de gekozen basisprocessen analyseren, synthetiseren en evalueren niet specifiek zijn voor ontwerpprocessen. Bijvoorbeeld ook de probleemoplossingsprocessen van specificatieproblemen zoals onderscheiden door Mettes en Pilot (1980) en van verklaringsproblemen zoals onderscheiden door Kramers-Pals (1989) zijn tot deze drie processen terug te brengen. Wèl heeft Teyken met deze drie processen de kern van het probleemoplossen getroffen. Het kenmerkende van ontwerpen is inderdaad dat de volgorde van de processen sterk kan variëren in tegenstelling tot b.v. specificatieproblemen. In de praktijk blijkt er echter toch een bepaalde volgorde te worden aangehouden. Een volgorde is ook aanwezig in het fasenmodel van Teyken: eerst oriënteren, vervolgens experimenteren en pas daarna uitvoeren. Een zeer interessant aspect aan dit fasenmodel is dat per gespecificeerde ontwerphandeling een duidelijk omschreven tussenprodukt is aangegeven. Als later het model wordt gebruikt als een interpretatiemodel voor een kwantitatieve en kwalitatieve analyse van protocollen kunnen deze tussenprodukten uitstekend dienst doen als afhankelijke variabelen voor het procesverloop.

In het *derde hoofdstuk* wordt onderzoek naar het leren probleemoplossen besproken. Teyken komt tot drie conclusies: a) voor de ontwikkeling van wendbaar oplossingsgedrag is het van belang dat oplossingsactiviteiten en vaardigheden zowel op een algemeen als een meer specifiek niveau onderwerp van oriëntatie en reflexie zijn; b) het leren oplossen van complexe problemen veronderstelt dat de probleemoplosser beschikt over een interne representatie van het betreffende oplossingsproces; en c) een voortgezette ontwikkeling van Gal'perins onderwijsleertheorie vereist onderzoek naar probleemoplossen dat een onvolledige oriënteringsbasis vereist en naar de wijze waarop deze kan worden uitgebreid.

Onderwijs in het leren probleemoplossen is volgens Teyken produktief van karakter. Dat wil zeggen, dat het gericht is op het leren toepassen van vakinhoudelijke kennis en het doelgerichter, systematischer en bewuster leren aanpakken van problemen. Om deze doelen te realiseren worden leerlingen/studenten geïnformeerd over de oplossingsactiviteiten en vaardigheden of werkwijzen die voor het oplossen van een bepaalde categorie van problemen moeten worden gebruikt. De verscheidenheid aan curricula en werkwijzen in de literatuur wordt vervolgens samenvattend besproken aan de hand van een classificatie met twee dimensies: a) niveau van de leerresultaten (procesniveau; deelprocesniveau, en niveau van oplossingsactiviteiten en vaardigheden); en b) moment van verwerking van de handelingsvoorschriften (vooraf, tussentijds, en achteraf).

De onderwijsleertheoretische principes zijn voornamelijk geïnspireerd door Oosteuropese theoretici als Gal'perin, Lompscher, etc. Terecht wordt geconstateerd dat de klassieke benadering van Gal'perin – het vooraf verstrekken van een volledige oriënteringsbasis – niet toepasbaar is voor het leren ontwerpen. Er moet met een onvolledige oriënteringsbasis worden gewerkt die met behulp van reflectie na afloop van het handelen wordt verbeterd (vgl. ook Terlouw, 1987). Wèl mis ik hier een voor de hand liggende relatie met het ervaringsleren (zie b.v. Erkamp, 1980). Een probleem vind ik voorts ook dat uit de gekozen onderwijsleertheorieën en het literatuur-overzicht van onderwijs in probleemoplossen, geen duidelijke richtlijnen voor onderwijsontwikkeling worden geformuleerd. Naar mijn mening had hier een concept als 'onderwijsfuncties' goede diensten kunnen bewijzen (Mettes & Pilot, 1980; Smuling, e.a., 1982; Terlouw, 1987). Onderwijsfuncties geven in algemene bewoordingen aan welke activiteiten in het onderwijs moeten worden ondernomen om de gewenste leerprocessen te stimuleren.

In het *vierde hoofdstuk* wordt een beschrijving gegeven van het bestaande en experimentele curriculum. De directe aanleiding tot de onderwijsontwikkeling en het onderzoek vormde het verzoek van de vaksectie om de eerste jaars cursus ontwerpen te reviseren op grond van onderwijskundige inzichten in het algemeen en leren probleemoplossen in het bijzonder. Op grond van gesprekken met docenten en studenten werd een globale beschrijving van het bestaande programma opgesteld en werden vier tekortkomingen bepaald: a) procesmatig werken als te globaal principe; b) gebrekkige veruiterlijking van ontwerpprocessen; c) sterk individueel gerichte onderwijsleerprocessen; en d) een te open karakter van de aangeboden ontwerpproblemen. Deze tekortkomingen veroorzaakten problemen in de begeleiding en de beoordeling. Het bestaande curriculum richt zich vooral op de experimentatie- en uitvoeringsprocessen. In het experimentele curriculum ligt met name de nadruk op het plannen en uitvoeren van oriëntatieprocessen en het reflecteren daarop. De kern daarvan is het laten gebruiken van een *planningsheuristiek*. Dit onderwijsinstrument bestaat uit tien oriënteringskaarten die de belangrijkste ontwerpactiviteiten en vaardigheden betreffen. Elke oriënteringskaart omvat meerdere handelingsvoorschriften in de vorm van als-dan regels die de uitvoering van de betreffende activiteit/handeling stuurt. In een bijlage is het experimentele curriculum beschreven.

De planningsheuristiek (blz. 173 t/m 176) in de vorm van tien oriënteringskaarten is zonder meer een vondst. Niet alleen is dit instrument een concrete handreiking voor de onderwijspraktijk – mede gelet op heldere gebruiksprocedure (blz. 187) –, het kan echter ook goed worden gebruikt als een instrument om oriëntatieprocessen in het probleem-

oplossen te onderzoeken. De beschrijving in de bijlage van het experimentele curriculum is wel behoorlijk gedetailleerd, echter, niet erg strak. Het kost enig zoekwerk om er achter te komen dat het om negen bijeenkomsten gaat. Ook de tijdsomvang staat niet gespecificeerd. Voorts lijkt me de relatie met de eerder geformuleerde onderwijsleertheoretische uitgangspunten nogal losjes; het precieze verband wordt nergens geëxpliciteerd. Een ordening van de specifieke onderwijsmiddelen en -procedures aan de hand van de eerder genoemde onderwijsfuncties had hier wellicht meer helderheid verschaft.

In het *vijfde hoofdstuk* wordt de opzet van het evaluatieonderzoek beschreven. Twee vraagstellingen staan centraal: 1. Leidt het experimentele curriculum (i.h.b. de genoemde planningsheuristiek) tot een verbetering van de ontwerpprocessen? en 2. Resulteert een verbetering van de ontwerpprocessen in een verbetering van de ontwerpprodukten? Er wordt gebruik gemaakt van een variant van een quasi-experimenteel pretest-posttest control group design, namelijk één met herhaalde metingen. De evaluatie wordt opgezet vanuit de doelstellingen van het experimentele curriculum. Op ontwerpproces-niveau worden de kennis van als-dan ontwerpregels, de frequenties en overgangsfrequenties van vaardigheden en de bestede tijd aan de oriëntatie geëvalueerd. Op ontwerpproduktniveau worden tussenprodukten en eindprodukten geëvalueerd. Tussenprodukten betreffen het aantal ideeën/oplossingen, het aantal representaties, het aantal deelproblemen en de procesplanning. De eindprodukten worden op beeldende, technische en functionele dimensies paarsgewijs beoordeeld.

Het onderzoek vindt zijn zwaartepunt in de afname van individueel gerichte ontwerpsessies waarin de proefpersonen in de voor- en nameting al hardopdenkend en schetsend een verpakking moeten ontwikkelen voor een aantal objecten. Er namen 16 willekeurig geselecteerde eerstejaars studenten van de NLO Ubbo Emmius (sectie handvaardigheid) deel: 8 studenten met de vakkencombinatie handvaardigheid/tekenen in de experimentele groep; 8 studenten met de vakkencombinatie handvaardigheid/textiele werkvormen in de controlegroep. In het hoofdstuk wordt niet aangegeven of de onderzoeksopzet een formatief of een summatief doel heeft. Pas in hoofdstuk 10 wordt in de inleiding gesteld dat het om een formatief gericht evaluatieonderzoek ging met als zwaartepunt de descriptie en analyse van de ontwerpprocessen van de experimentele groep (blz. 122). Als we dit echter vergelijken met de vraagstellingen in hoofdstuk 5 (blz. 41), dan wordt niet duidelijk wat nu voor de onderwijsontwikkeling het formatieve karakter is. Bovendien vind ik het ook merkwaardig dat in deze formatieve evaluatie niet de onderwijsmiddelen en -procedures worden meegenomen. Ook al gelet op het feit dat er sprake is van een quasi-experimenteel design waarvoor het voor een goede interpretatie van de onderzoeksresultaten noodzakelijk is dat de uitvoering van de experimentele variable (het experimentele curriculum) wordt gecontroleerd. Pas in het replicatie-onderzoek wordt de functierealisatie van de onderwijsmiddelen geëvalueerd. Overigens blijken er na de uitvoering van het onderhavige onderzoek toch wel belangrijke verbeteringen te zijn aangebracht in het experimentele curriculum (zie blz. 123). Op grond van welke evaluatieve informatie dit is gebeurd, blijft duister.

Hoofdstuk 6 behandelt de constructie van de onderzoeksinstrumenten. Er wordt op vier onderwerpen ingegaan: de ontwikkeling van een aanvangs- en criteriumprobleem, van de proceskennistoets, van het transcriptiesysteem voor de hardopdenkprotocolanalyse en de bepaling van de betrouwbaarheid en validiteit van de produktbeoordeling. De keuze van een equivalent aanvangs- en criteriumprobleem leverde problemen op, omdat er in het vakgebied geen concensus bestond over een probleemclassificatie naar moeilijkheidsgraad. Uiteindelijk is door de onderzoeker zelf een klassificatie gemaakt op basis van drie variabelen: kenmerken van de probleemoplosser, van het ontwerpprocesniveau en van het produkt. De proceskennistoets was een aanvultoets naar aanleiding van een opdracht en betrof de kennis van als-dan regels zoals die waren geoefend met de planningsheuristiek. Een transcriptie-systeem werd ontwikkeld ter analyse en codering van de hardopdenkprotocollen. De grondstructuur van het systeem wordt bepaald door vier gedragscategorieën in de uitspraken: de tijdsoriëntatie (retrospectief, actueel, prospectief); het constaterende of oordelende karakter; betreffende het proces of het produkt; en kenmerken van het proces/produkt. De betrouwbaarheid en validiteit van de produktbeoordelingen werden per ontwerpdimensie nagegaan met behulp van respectievelijk Cronbachs alpha en factoranalyse. De betrouwbaarheid is veelal acceptabel (rond de .72); de validiteit is wat minder.

Terecht wordt de afwezigheid van een algemeen aanvaarde probleemclassificatie naar moeilijkheidsgraad naar voren gehaald als een belangrijke belemmering bij het onderzoek naar complexe probleemoplossingsprocessen. Naast de genoemde variabelenbenadering wordt het aantal onderkende deelproblemen ook als en mogelijke maatstaf genoemd. Op zichzelf zijn deze classificatiemogelijkheden vanuit analytisch oogpunt een verrijking. Een probleem is evenwel dat juist complexe problemen zich kenmerken door componenten die analytisch niet zo netjes van elkaar zijn te onderscheiden. Onafhankelijke deelproblemen bijvoorbeeld zijn maar zeer ten dele te onderscheiden. Dergelijke problemen heten niet voor niets 'slecht gedefinieerd'. Mijns inziens is het een meer vruchtbare weg problemen te classificeren op basis van de gewenste abstractieniveaus van de (im)materiële handelingen

die nodig zijn om tot een bevredigende oplossing van het probleem te komen (zie voor de abstractieniveaus Van Parreren, 1979; Vos, 1989). De variabelen- en deelproblemen-benadering van Teyken zijn vermoedelijk wel goede uitgangspunten om dergelijke handelingen te formuleren.

In *hoofdstuk 7* worden de resultaten van de kwantitatieve analyse van oriëntatietijden, als-dan regels en (tussen)produkten gerapporteerd. Het blijkt dat de experimentele groep het op alle variabelen op procesniveau – kennis van als-dan regels, aantal geplande stappen in het ontwerpproces; aantal representaties, de duur van de oriëntatieprocessen, produktie van ideeën en oplossingen, en het aantal onderscheiden deelproblemen – beter doet dan de controlegroep. Echter, ondanks al deze veranderingen op procesniveau werden tussen de totaalscores van de eindprodukten van beide groepen géén significante verschillen gevonden. Alleen de scores op de ontwerpdimensie gaven op 10% niveau een verschil te zien ten gunste van de experimentele groep. Kortom: acceptatie van de proceshypothese, maar verwerping van de produkthypothese.

Deze discrepantie tussen proces- en produktgegevens is zeer interessant. De auteur verdient dan ook alle lof dat hij deze discrepantie zo consciëntieus rapporteert en probeert naar verklaringen te zoeken in de volgende hoofdstukken. In het kader van onderwijsontwikkeling zijn dergelijke discrepanties tussen het gewenste leereffect (produkt) en tussenresultaten van ingezette onderwijsmiddelen (proces) – in dit geval de planningsheuristiek – zeker bij een eerste prototype heel vaak aan de orde. Deze omstandigheid is een gevolg van de stand van wetenschappelijke kennis in de onderwijskunde: we weten niet zo precies welke onderwijsmiddelen onder welke condities tot welke leereffecten leiden. Bijgevolg is er veelal sprake van een zo goed mogelijk, theoretisch onderbouwde gissing, op grond waarvan onderwijs wordt ontwikkeld, uitgevoerd en geëvalueerd. Komen de verwachte leereffecten niet voor, dan moet er in eerste instantie binnen de gekozen theorie die aan het onderwijsprogramma ten grondslag ligt, naar verklaringen worden gezocht. Het gebruik van een onderwijsleertheorie is derhalve cruciaal. Teyken volgt wel min of meer deze weg, echter te veel blijft impliciet. Juist op het punt van de theoretische onderbouwing gaat het hierboven al gesignaleerde losse verband tussen onderwijsleertheorie en experimenteel onderwijsprogramma zich wreken. Ik kom daar straks nog op terug bij de verklaringen die Teyken geeft voor de discrepantie.

Hoofdstuk 8 bespreekt de resultaten van een kwantitatieve analyse van de hardopdenkprotocollen. Uit de analyse van de frequenties en de overgangsfrequenties blijkt dat er aanzienlijke verschillen bestaan tussen het aanvankelijk en het gevorderd ontwerpgedrag, hoewel de grondstructuur van het gedrag niet verandert. In de grondstructuur blijven twee met elkaar verbonden substructuren aanwezig: een doel- en een middelstructuur. De doelstructuur bestaat uit prospectieve analyse- en evaluatieuitspraken met betrekking tot het eindprodukt of tussenprodukten alsmede ideeënproducerende uitspraken. De middelstructuur betreft handelingen aangaande toetsing en uitvoering van ideeën en oplossingen. In het gevorderd ontwerpgedrag blijkt de middelstructuur en de samenhang tussen doel- en middelstructuur te zijn versterkt. Het gevorderde gedrag wordt gekenmerkt door een sterke actuele gerichtheid, een verscherping van de produktgerichtheid en de aanwezigheid van reflectie op het proces naar aanleiding van fouten. Een spilfunctie vervult het analyseren en beoordelen van tussenprodukten. De verwachte vergroting van doelmatigheid, beheersing en bewustheid van het ontwerpgedrag wordt aangetroffen. Opmerkelijk genoeg is dat niet het geval met de verwachte vergroting van de planmatigheid van het gedrag.

Het laatste is wellicht te verklaren uit het feit dat een grotere bewustheid, beheersing, e.d. van het ontwerpgedrag, de behoefte aan planning van dit gedrag doet afnemen. Waarschijnlijk is in de beheersing de planning impliciet meegenomen. Terlouw en Mettes (1986) vonden b.v. in een onderzoek naar het beleidsontwerpen bij experts en studenten dat de laatsten veel meer planden dan de eersten. Bijna de helft van het handelen van de studenten betrof planning.

In *hoofdstuk 9* komen de resultaten van de kwalitatieve analyse van de hardopdenkprotocollen aan de orde. De kwalitatieve analyse is gericht op gedragsaspecten die in het oorspronkelijke fasenmodel (zie hoofdstuk 2) werden verwoord. Deze hebben betrekking op de voortgebrachte tussenprodukten en op het verloop van de ontwerpactiviteiten en de vaardigheden. De auteur beschrijft aan de hand van deze gedragsaspecten, te onderkennen aan tussenprodukten en ontwerpactiviteiten, op gedetailleerde wijze de inhoudelijke verschillen tussen aanvankelijk en gevorderd ontwerpgedrag (overzicht figuur 9.2. blz. 111/112). In de interpretatie van deze resultaten komt hij tot de volgende conclusies:

- De gegevens bevestigen de bestaande theoretische inzichten aangaande probleemoplossen. Oplossingsprocessen verlopen van algemeen naar specifiek. De kern van de oriëntatieprocessen bestaat uit de ontwikkeling van een ontwerp. De kern van de uitvoeringsprocessen bestaat uit het materialiseren van het ontwerp.
- Ontwerpproblemen zijn complexe problemen. In het aanvankelijk ontwerpgedrag is men zich van deze complexiteit niet bewust. In het gevorderd ontwerpgedrag komt deze complexiteit wèl tot bewustzijn. Men realiseert zich de omvang van de ontwerpruimte. In dit kader worden vier competentieniveaus onderscheiden.

- De gegevens bevestigen eveneens het ook in andere domeinen gevonden verschil tussen aanvankelijk en gevorderd gedrag. Het aanvankelijk gedrag verloopt ad hoc-achtig, terwijl aan het gevorderd gedrag een handelingsplan ten grondslag ligt. Teyken komt tot de conclusie dat de proefpersonen in de experimentele groep op reproduktief niveau over een probleemaanpak beschikken die het gedrag stuurt en controleert. Hij ziet dit als een direct leerresultaat van het experimentele curriculum waarin immers het werken met een planningsheuristiek centraal stond.
- Teyken oppert vier verklaringen voor de discrepantie tussen proces- en produktgegevens, door hem aangeduid als het flessehalsfenomeen. Voor drie verklaringen vindt hij enige ondersteuning in de gegevens: het ontbreken van voldoende technische kennis en geoefendheid, en een tekort aan beeldende (functionele) kennis en ervaring. Deze hiaten zouden er toe leiden dat wel meer maar niet de kwalitatief betere oplossingen voor het ontwerpprobleem worden onderkend en gekozen.

De gevonden resultaten aangaande de probleemaanpak sluiten aan bij de resultaten die ik heb gevonden in het onderzoek om studenten in twee experimentele cursussen een systematische probleemaanpak voor beleidsontwerpen te leren (Terlouw, 1987). Ook daarin bleken de studenten zich een systematiek eigen te hebben gemaakt die het gedrag stuurde en controleerde. Bovendien bleek dat een relatief groot aantal studenten de systematiek zelf ook als een object van verandering zag. De planningsheuristiek bestond in het beleidsontwerpgeval uit een checklist, die, in tegenstelling tot wat Teyken onmogelijk achtte, uit een gewenst handelingsverloop was afgeleid. Hij ging er ten onrechte vanuit dat studenten in de Twentse probleemoplos-cursussen een gewenst handelingsverloop (GHV) krijgen voorgelegd. Dit is niet het geval. Een gewenst handelingsverloop (GHV) is een doelstellingsformulering waarin *voor een onderwijsontwerper* de uit te voeren handelingen en de volgorde daarvan rationeel zijn gereconstrueerd. Een leermiddel voor studenten dat daaruit wordt afgeleid – een systematische probleemaanpak (SPA) – moet vooral het bedoelde leerproces van studenten ondersteunen. Voor natuurkundige problemen is het mogelijk een gewenste volgorde aan te geven; voor ontwerpproblemen is dat niet mogelijk.

Hoofdstuk 10 beschrijft de opzet van het replicatieonderzoek. Met het replicatieonderzoek wordt nagegaan of de gevonden leereffecten in de vorm van wijzigingen in het ontwerpgedrag systematisch van aard zijn of op toeval berusten. Ongemerkt zij,rdat op een aantal punten het experimentele curriculum is veranderd: in het begin toespitsing op de introductie van de planningsheuristiek; inperkingen, toevoegingen en specificaties van de planningsheuristiek. Het replicatieonderzoek bestaat uit drie deelonderzoeken: naar ontwerpprocessen, naar ontwerpprodukten en naar de mate van realisatie van de geplande onderwijsleerprocessen. Ten aanzien van ontwerpprocessen worden observatiegegevens verzameld van een aantal tussenprodukten (b.v. het aantal geplande ontwerpactiviteiten; de duur van het oriëntatieproces; het aantal geformuleerde deelproblemen; etc.). Voor de ontwerpprodukten worden de oordelen verzameld van vijf onafhankelijke beoordelaars die op een 9-puntsschaal de eindprodukten beoordelen op criteria aangaande de beeldende, functionele en technische dimensie. Het deelonderzoek betreffende de evaluatie van de onderwijsleerprocessen heeft twee delen: a) onafhankelijke observatoren gaan de mate van relalisatie na van de geplande onderwijsleerprocessen; b) deelnemers van het experimentele onderwijs beoordelen op een uitsprakenlijst in het bijzonder de planningsheuristiek in zijn functie als hulpmiddel bij het ontwerpen.

De onderzoeksopzet betreft een quasi-experimenteel experimental control group design met herhaalde metingen. Bij de instrumenten is de betrouwbaarheid (= interne consistentie) nagegaan en zijn (Cronbachs) alpha's gevonden tussen de .68 en .82.

In *hoofdstuk 11* worden de resultaten van het replicatieonderzoek en de interpretatie gerapporteerd. Dezelfde resultaten als in de hoofdstukken 7, 8 en 9 worden gevonden: wèl significante verschillen tussen experimentele en controlegroep aangaande de ontwerpprocessen, niet aangaande de eindoordelen voor de eindprodukten. De significante verschillen blijken zich vooral te concentreren in de inrichting van het oriëntatieproces en in de hoeveelheid tijd die aan dit proces wordt besteed. Voorts wordt met enige voorzichtigheid geconcludeerd dat het experimentele curriculum differentiële leereffecten oproept.

Teyken ziet de volgende verklaring als acceptabel voor het feit dat de kwalitatief betere ontwerpprocessen niet resulteren in betere eindprodukten: '. . .doordat men op grond van technische overwegingen besluit om een kwalitatief mindere oplossing uit te voeren en/of omdat men door het ontbreken van beeldende kennis de betere ideeën en oplossingen niet onderkent en/of geen betere ideeën en oplossingen produceert' (blz. 134). Op zichzelf is deze verklaring wel plausibel, echter, het blijft onbevredigend dat het onduidelijk is waarom deze verklaring nu als het meest relevant wordt gezien. Een methodisch betere weg was m.i. geweest terug te keren naar de gebruikte onderwijsleertheorie: Gal'perin. Als we er vanuit gaan dat het oriënteren in deze onderwijsleertheorie centraal staat (Van Parreren, 1989), dan zou in eerste instantie de verklaring moeten worden gezocht in een tekortschietende oriëntering in het onderwijs, in het bijzonder in de planningsheuristiek. Daarvoor hoeft men het idee van de onvolledige oriënteringsbasis nog niet te laten vallen. In deze oriënteringsbasis ontbreekt blijk-

baar iets cruciaals waardoor studenten zich niet optimaal voorbereiden op het gewenste produkt. In Teykens verklaring zit inderdaad een belangrijke voorwaarde: beschikking over relevante technische en beeldende kennis. De planningsheuristiek beperkt zich tot de strategische component van het probleemoplossend handelen. Uit ander onderzoek is bovendien bekend dat het juist moet gaan om de integratie van deze strategische component en relevante, goed beschikbare domeinkennis (zie b.v. Mencinskaja, 1966; Van Parreren, 1974; Mettes & Pilot, 1980; Elshout, 1985 en recent De Corte, 1989). Als de koppeling tussen onderwijsleertheorie en experimenteel onderwijsprogramma steviger was geweest (b.v. aan de hand van onderwijsfuncties), had er wellicht een preciezer herontwerp van de experimentele cursus kunnen plaatsvinden op grond van de evaluatie-resultaten. De kans op consistentie tussen proces en produkt was daarmee toegenomen. Gelet op de complexiteit van het onderwerp zou mij het echter niet hebben verbaasd als de discrepantie nog niet was weggewerkt. Ik vermoed wel dat de goede weg is ingeslagen. Het zou nog wel enkele ontwikkelcycli hebben gekost om helemaal tevreden te zijn.

In *hoofdstuk 12* wordt besloten met een slotbeschouwing waarin a) de onderzoeksuitkomsten worden vergeleken met verwant Nederlands onderzoek; en b) een descriptief model voor de structuur van ontwerpprocessen wordt gepresenteerd.

Een aantal Nederlandse onderzoekingen blijkt op produktniveau (en soms ook procesniveau) wèl significante resultaten gevonden te hebben (Mettes & Pilot, 1980; Been e.a., 1985; Brokken & Been, 1985). Dit zou te verklaren kunnen zijn uit het type problemen en uit de aard van de vakinhoudelijke kennis die aan de orde zijn. In alle gevallen betreft het natuurwetenschappelijke, exacte problemen waarvoor maar één goede oplossing bestaat en waarvoor de benodigde kennis duidelijk omschreven en 'eenvoudig toegankelijk' is. Kortom, een (nagenoeg) volledige oriënteringsbasis die op heldere wijze toegankelijk kan worden gemaakt. Teyken trekt uiteindelijk de volgende conclusies:

- ontwerpprocessen kunnen met behulp van een onvolledige oriënteringsbasis worden verbeterd;
- door reflexie op de doorlopen processen kan een onvolledige oriënteringsbasis worden uitgebreid en verfijnd;
- óók voor het ontwerpen op het terrein van grafische en industriële vormgeving is het noodzakelijk dat vakinhoudelijke kennis en kunde op het terrein van oplossingsmethoden beschikbaar is. In dit kader gaat het dan vooral om beeldende/functionele kennis en technische kennis en geoefendheid.

In het descriptieve model voor de structuur van ontwerpprocessen in het oplossingsgedrag worden drie hiërarchische niveaus onderscheiden: deelprocessen, ontwerpprocessen, en vaardigheden of operaties. Elk niveau heeft z'n eigen tussenprodukten. Deze tweedeling – oplossingsgedrag en tussenprodukten – wordt gekruist met drie typen criteriumvariabelen – extensieve (frequentie), intensieve (meer meetmomenten) en successieve variabelen – en resulteert in een categorisering van criteriumvariabelen voor ontwerpprocessen op produktief niveau. Dit kan worden gebruikt voor de analyse van hardopdenkprotocollen.

Ten slotte breekt de auteur nog een lans voor onderzoek naar handelingsvoorschriften in de vorm van als-dan regels op de verschillende niveaus van het descriptieve model. Het betreft descriptief onderzoek naar overeenkomsten en verschillen op verschillende expertise-niveaus van ontwerpers van verschillende origine. Als methode beveelt Teyken de hardopdenk-methode, gekoppeld met de interviewmethode aan.

Ik kan het met de meeste dingen die de auteur in dit hoofdstuk stelt, eens zijn. Met name het opgestelde descriptieve model voor de structuur van ontwerpprocessen in het oplossingsgedrag is een goede bijdrage om hardopdenkprotocollen uit verschillende ontwerpdomeinen op dezelfde manier te analyseren.

Het geheel overziende is mijn uiteindelijk oordeel over het werk positief. Teyken heeft niet alleen voor onderzoekers en onderwijsontwikkelaars aangaande het beeldend ontwerpen een belangwekkend onderzoek uitgevoerd. Oók voor andere ontwerpdomeinen en voor probleemoplossen in het algemeen, heeft hij een aantal vensters geopend of tenminste aangewezen, waardoor wij een stukje verder kunnen komen met onderzoek èn onderwijs. Dergelijke onderzoekers/onderwijzers zijn zeldzaam in onze discipline. Des te meer moeten wij het betreuren dat hij niet meer onder ons is.

C. Terlouw

Literatuur

Been, P.H., e.a., *Leren oplossen van praktische farmaco-therapeutische problemen: effect van een computergestuurde cursus.* Groningen: RUG/COWO, 1985.

Boer, S.J. de, *Decision methods and techniques in methodical engineering design.* Proefschrift Universiteit Twente. Enschede: UT/WB, 1989.

Brokken, F.B. & P.H. Been, *Leren van expertgedrag bij het kiezen van statistische toetsen: effecten van instructie en zelfontdekkend leren.* Groningen: RUG Instituut voor Onderwijskunde/COWO: 1985.

Corte, E. De, *Ontwerpen van krachtige onderwijsleeromgevingen.* Paper ORD'89, Leiden, mei 1989.

Elshout, J. J., *Problem Solving and education*. State of the art. Paper 1st EARLI-conferentie Leuven, 1985.

Erkamp, A. D., *Ervaringsleren; praktijkinformatie voor vorming en onderwijs*. Amersfoort: Studiecentrum NCVO, 1980.

Kramers-Pals, H., *Feitelijk handelingsverloop bij verklaringsproblemen in het chemie-onderwijs*. Paper ORD'89, Leiden, mei 1989.

Mencinskaja, N. A., *Développement de la pensée au cours du processus d'enseignement scolaire*. Recherche psychologiques de URSS. Moskou: 1966.

Mettes, C. T. C. W. & A. Pilot, *Over het leren oplossen van natuurwetenschappelijke problemen*. Dissertatie Universiteit Twente. Enschede: UT/OC, 1980.

Parreren, C. F. van, Het functioneren van leerresultaten. In: C. F. van Parreren & J. Peeck (red.), *Leerpsychologie en onderwijs 1*. Groningen: Tjeenk Willink, 1974.

Parreren, C. F. van, Niveaus in de ontwikkeling van het abstraheren. In: J. de Wit, e.a. (red.), *Psychologen over het kind*, deel 6. Groningen: Wolters-Noordhoff, 1979.

Parreren, C. F. van, *Ontwikkelend onderwijs*. Leuven: Acco, 1989.

Smuling, e.a. (red.), *Oriëntatie op leren en onderwijs*. Utrecht: Het Spectrum, 1982. (Inmiddels Wolters-Noordhoff).

Terlouw, C. & C. T. C. W. Mettes, De validiteit voor een systematische probleemaanpak voor het ontwerpen van beleid. *Tijdschrift voor Onderwijsresearch*, 1986, *11*, 203-216.

Terlouw, C., *De FUNDES-procedure in onderwijsontwikkeling*. Evaluatie van een procedure van onderwijsontwikkeling voor het leren probleemoplossen. Dissertatie Universiteit Twente. Enschede: UT/OC, 1987.

Vos, H., *Naar een oriënteringsbasis voor probleemoplossen in het vak Netwerkanalyse*. Paper ORD'89, Leiden, mei 1989.

Evaluatie van een beleidsmaatregel
Uitvoering en effecten van de nascholing zorgverbreding lezen

K. M. STOKKING en F. J. LEENDERS
Rijksuniversiteit Utrecht

Samenvatting

Vanaf 1984 wordt in het Nederlandse basisonderwijs een gericht zorgverbredingsbeleid nagestreefd. De eerste concrete beleidsmaatregel was een landelijke nascholingsoperatie, 'speerpunt lezen' (1985-1988). De opzet, uitvoering en resultaten daarvan zijn onderzocht. Dit artikel biedt eerst een korte beschrijving van het beleid, de opzet van het onderzoek en de in de literatuur beschikbare kennis over leesonderwijs, zorgverbreding en nascholing. Daarna bespreken we de beginsituatie op de basisscholen, de uitvoering van en deelname aan de nascholingsoperatie, en de resultaten daarvan. Deze laatste vallen tegen: Van de nascholingsoperatie zijn weinig specifieke effecten vastgesteld.

1 *Inleiding*

In 1984 is besloten tot een beleid, gericht op zorgverbreding in het basisonderwijs. Het voornaamste beleidsinstrument werd een landelijke nascholingsoperatie. In dit artikel wordt verslag gedaan van een onderzoek naar de uitvoering en resultaten daarvan. We schetsen kort het beleid (par. 2) en de opzet van het evaluatieonderzoek (3). Daarna volgt een samenvatting van de uitkomsten van een literatuurstudie (4) en van gesprekken met het Ministerie van O&W en de LPC over relevante evaluatiecriteria (5). Dan wordt de opzet van het uitgevoerde surveyonderzoek beschreven (6). Ten slotte worden de uitkomsten daarvan gepresenteerd, in drie stappen: de stand van zaken in het basisonderwijs voorafgaande aan de nascholing (7.1), de uitvoering van het beleid (7.2), en de resultaten daarvan (7.3). Het artikel besluit met enkele conclusies (8).

2 *Het beleid*

In 1984 constateerde de ARBO dat niet alle leerlingen de doelen van het basisonderwijs bereiken. Er treden vooral stagnaties op bij het leren lezen; deze zijn bovendien vaak aanleiding tot zittenblijven en verwijzing naar het speciaal onderwijs. Er werd besloten tot een meerjarig zorgverbredingsbeleid (Ministerie van O&W, 1984b), met als eerste speerpunt lezen (in de jaren '85-'88; bij een tweede speerpunt dacht men aan rekenen, in de jaren '88-'91) en als doelen: kwalitatief verantwoord basisonderwijs aan leerlingen met hardnekkige problemen en minder verwijzingen naar het speciaal onderwijs. Als primair aangrijpingspunt werd gekozen voor vergroting van de bekwaamheid van leerkrachten en als voornaamste middel daartoe voor een landelijke nascholingsoperatie. De PABO's kregen tot taak om jaarlijks een cursus van 32 uur te geven. De deelnemende scholen kregen extra formatie. De LPC dienden de benodigde materialen te ontwikkelen en de PABO-docenten te ondersteunen ('voorscholing'). De OBD's werden geacht de deelnemende schoolteams en leerkrachten te begeleiden ('nazorg'). Gedurende 3 schooljaren ('85-'88) konden budgettair gezien jaarlijks maximaal 22.500 cursisten meedoen, zodat in drie jaar alle leerkrachten basisonderwijs de cursus konden volgen. De operatie is in 1985 gestart met een cursus 'voorbereidend, aanvankelijk en technisch lezen' (cursus A). Vanaf januari 1987 werd ook 'begrijpend lezen' aangeboden (B), vanaf januari 1988 ook 'taalvaardigheden' (C).

3 *De opzet van het evaluatieonderzoek*

De nascholing 'speerpunt lezen' is onderzocht in opdracht van het Ministerie van O&W en met subsidie van de SVO. Het onderzoek richtte zich in het bijzonder op cursus A. Zie voor een uitgebreid verslag Stokking en Leenders (1988).

Bij de opzet van een evaluatieonderzoek moeten keuzen worden gemaakt. De afgelo-

pen 20 jaar hebben zich internationaal rondom evaluatieonderzoek diverse opvattingen gevormd, zoals: evaluatieonderzoek moet nagaan in hoeverre de doelen van een programma worden bereikt, moet informatie aandragen ter ondersteuning van beslissingen resp. moet nagaan welke effecten met een programma worden bereikt en waaraan die kunnen worden toegeschreven (zie o.a. Cronbach, 1982; Glass & Ellet, 1980; Van de Grift, 1984; De Groot, 1986; Hofstee, 1980; House, 1977; Stokking, 1979, 1982, 1984; Stufflebeam, 1972). In dit onderzoek is op verzoek van de opdrachtgever het accent gelegd op het meten van effecten van de nascholingsoperatie, uitgaande van de doelen die door het beleid zijn geformuleerd en met het oog op besluitvorming over een volgend speerpunt binnen het zorgverbredingsbeleid. De keuze van de criteria en overige variabelen dienden te worden afgestemd op de gestelde *doelen* en *uitvoeringscondities* (zie par. 2 en par. 5) en op de beschikbare *kennis* (zie par. 4). Het onderzoek behelsde de volgende onderdelen:

a. studie van beleidsdocumenten, vakpublikaties en onderzoeksliteratuur over leesonderwijs, zorgverbreding en nascholing;
b. inventarisatie van verwachtingen en evaluatiecriteria m.b.t. de cursus bij voorbereiders en uitvoerders van het beleid;
c. surveyonderzoek bij PABO's, OBD's en basisscholen;
d. beoordeling van de landelijk ontwikkelde materialen door deskundigen t.a.v. leesonderwijs en leesproblemen;
e. case studies bij enkele PABO's, docenten en scholen, via interviews en observaties (cursusbijeenkomsten, leeslessen);
f. koppeling van de surveygegevens aan toetsgegevens van het CITO, bij scholen die aan beide onderzoeken deelnamen.

In dit artikel wordt gerapporteerd over a. (par. 4), b. (par. 5) en c. (par. 6 en par. 7). De uitkomsten van a. en b. leidden tot de conclusie dat grote verschillen konden worden verwacht in de beginsituatie op de scholen en in de uitvoering van de cursussen. Deze variatie kon alleen voldoende recht worden gedaan via een grootschalig surveyonderzoek. Dit bracht met zich mee dat onderdeel c. het hoofdbestanddeel ging vormen van het evaluatieonderzoek. De onderdelen d., e. en f. werden kleinschalige en kwalitatieve aanvullingen op het surveyonderzoek. De uitkomsten daarvan bevestigden de uitkomsten van het surveyonderzoek (zie Stokking & Leenders, 1988). Om deze reden en omwille van de hier beschikbare ruimte laten we deze onderdelen hier verder buiten beschouwing.

De literatuurstudie (onderdeel a., zie par. 4) leidde niet alleen tot de keuze voor een grootschalige opzet maar diende tevens als bron voor de keuze en operationalisering van relevante variabelen (zie par. 6).

4 *De literatuur*

4.1 *Leesonderwijs*

Door diverse auteurs worden enigszins uiteenlopende percentages kinderen genoemd, die problemen hebben met het (leren) lezen. Deze verschillen hangen samen met de operationalisering van en de gehanteerde criteria bij het begrip 'leesproblemen'. Men let daarbij op verschillende tijdstippen (groep 3, groep 5, groep 8 basisschool) en verschillende aspecten/problemen (leestechniek, leesbegrip, leesbeleving). De schattingen variëren van 5-17%. (Van Dongen, 1984; Mommers, 1985a, 1985b; Mommers & Aarnoutse, 1984; SLO, 1986; Vinke, Reitsma, Van der Leij, 1985; Werts, 1985; Winnubst, 1985).

Doordat de beoordelingscriteria naar vaardigheid en tijdstip niet scherp zijn omschreven worden problemen door leerkrachten vaak pas geleidelijk onderkend. Veel leerkrachten gebruiken daartoe de voortgang van de rest van de klas als maatstaf. Kinderen met problemen raken vaak steeds verder achterop: achterstanden cumuleren. Het voorspellen van de ontwikkeling van het lezen en het vroegtijdig onderkennen van kinderen die problemen (zullen) ondervinden is echter slechts in beperkte mate mogelijk gebleken (Van Dongen, 1984, 1985; Melis, 1985; Van der Leij, in SVO 1985; Mommers, 1985a; Mommers, Aarnoutse, Smits, 1985; SLO, 1986; Smits, Mommers, Aarnoutse, 1985; Vinke et al. 1985). Na het aanvankelijk lezen gaan technisch en begrijpend lezen uit elkaar groeien: op een gegeven moment kan sprake zijn van correct verklanken zonder te begrijpen, en begrijpen zonder precies te lezen wat er staat. Voor vlot technisch lezen is snelle herkenning nodig. Zwakke lezers blijven echter veelal spellend of radend lezen. Ook het leunen op context-informatie is problematisch,

omdat daarbij mede de algemenere taalvaardigheid (woordenschat, basiskennis, intelligentie) in het geding is, waarin zwakke lezers meestal ook zwak zijn. (Van Dongen, 1984; Mommers et al., 1985; Roelofs, Visser, Boland, Mommers, 1987).

Over het voortgezet en begrijpend leren lezen is weinig bekend. Diverse auteurs bevelen aan om kinderen in ieder geval zoveel mogelijk te laten lezen (frequentie bevordert vaardigheid bevordert plezier bevordert frequentie). Problematisch is echter dat juist zwakke lezers vaak moeilijk tot lezen zijn te brengen, mede door het ontbreken van voldoende geschikte materialen (geringe vaardigheid en weinig interessante leesteksten remmen plezier en frequentie). (Van Laarhoven, 1985; Mommers, 1985a, 1985b; Reitsma 1983; SLO, 1986; Vernooy, 1985; Werts, 1985; Winnubst, 1985). Van diverse kanten is er kritiek op het gangbare voortgezet technisch lezen (teveel accent op correct verklanken, via het systeem van hardop-leesbeurten, waardoor geen rekening wordt gehouden met verschillen tussen kinderen, de hoeveelheid lezen onnodig wordt beperkt, en elk foutje wordt bestraft). Ook het begrijpend leesonderwijs lijkt uniform te worden gegeven, met weinig echte instructie (Van Dongen, 1985; Mommers & Aarnoutse, 1984; Werts 1985; Weterings & Aarnoutse, 1986).

Buitenlandse auteurs komen tot soortgelijke conclusies, ten aanzien van het vóórkomen van leesproblemen, het belang van gedifferentieerde materialen, het ontbreken van interessante leesteksten voor zwakke lezers, de weinig zinvolle aanpak van het voortgezet technisch lezen, het gebrek aan instructie bij begrijpend lezen, en de grote rol die de leerkracht te vervullen heeft bij de verbetering van het leesonderwijs en het opvangen van kinderen met leesmoeilijkheden (Alm, 1984; Carter, 1984; Chauveau, 1985; Pink & Leibert 1986). Leesonderwijs lijkt te vragen om variatie en differentiatie. Dit vergt keus uit materialen, oefeningen, teksten, en grote professionaliteit van de leerkracht. Deze zal grondig bekend moeten zijn met de theorie over het leren lezen en de in gebruik zijnde leesmethode, en alert moeten zijn op problemen en daarop flexibel kunnen inspringen (Van Dongen, 1984, 1985; Mommers, 1985b; Mommers & Aarnoutse, 1984; SLO, 1981, 1986).

4.2 *Zorgverbreding*

Aanleiding voor zorgverbreding vormen hardnekkige problemen van kinderen bij het verwerven van instrumentele vaardigheden (lezen, taal, rekenen) en sociaal-emotionele en gedragsproblemen (motivatie, concentratie). In het laatste jaar van het bestaan van de lagere school ('84-'85) werden door bevraagde leerkrachten bij 14% van de leerlingen leesproblemen genoemd, bij 16% rekenproblemen, bij 10% concentratie- en motivatieproblemen, bij 6% nervositeit en angsten en bij 7% hyperactiviteit en agressiviteit (Span, Abbring, Meijer, 1985). Een vrij groot deel van de problemen komt bij dezelfde leerlingen voor. Diverse auteurs stellen dat de categorisering van een leerling als één met problemen afhangt van percepties en beoordelingen van de leerkracht (Chauveau, 1985; Meijer, 1986; Vos, 1986). Als oorzaken van deze problemen worden door leerkrachten genoemd: concentratie- en motivatieproblemen bij de kinderen (Wolbert, Schaap-Hummel, Span, 1986). Wolbert c.s. noemt zelf als oorzaken echter ook differentiatieproblemen. Immers: kinderen in een groep maken zich met verschillende snelheden en resultaten leerstof eigen. Anderen wijzen op de (toegenomen) diversiteit in de samenstelling van de schoolbevolking en de geringe aandacht in het onderwijs voor het individuele kind (zie ook Meijer, 1982). Als gevolgen van de problemen worden genoemd: toenemende uitval, zittenblijven en verwijzing naar het speciaal onderwijs (Van Dongen, 1984).

Om zorgverbreding daadwerkelijk te realiseren worden veel aanbevelingen gedaan. Deze hebben zowel betrekking op het schoolniveau, bijv. het maken van afspraken binnen het team over continuïteit in de aanpak tussen de groepen, als het leerkrachtniveau, bijv. deskundigheidsbevordering. Een deel van de aanbevelingen heeft betrekking op algemenere maatregelen ten aanzien van klasseorganisatie, instructie, werkvormen. Meer specifiek op zorgverbreding gerichte aanbevelingen betreffen de attitude van leerkrachten tegenover kinderen met problemen, extra materialen, minimumdoelen, individuele leerwegen, foutenanalyse, leerlingbesprekingen, handelingsplannen. Als voorwaarden voor zorgverbreding worden vaak genoemd individualisering en differentiatie (ARBO, 1984; Meijer, Verwilligen-Prummel, Rippen, 1984;

Inspectie, 1985; Wolbert e.a., 1986). Over de feitelijke situatie met betrekking tot zorgverbreding is uit de literatuur niet veel bekend. Naar processen van verwijzing van leerlingen naar speciaal onderwijs wordt al enige jaren onderzoek gedaan. Daaruit komen indicaties naar voren met betrekking tot grote verschillen in pedagogisch-didactische aanpak tussen basisscholen onderling (zie Meijer, 1982; Meijer e.a., 1984). Volgens de Inspectie (1985) is er op scholen die herkenbaar iets aan zorgverbreding doen sprake van een brede interpretatie van dat begrip, maar blijven de activiteiten feitelijk beperkt tot de kinderen met problemen bij lezen en rekenen. Span e.a. (1985) komen tot de conclusie dat 13% van de leerlingen op de basisscholen de leerkrachten zorgen baren en op de school zelf moeten worden opgevangen. In de praktijk is dit primair een zaak van de individuele leerkracht. Reezigt, Van Dijk, Bosveld (1986), Span e.a. (1985) en Wolbert e.a. (1986) rapporteren over differentiatie bij het leesonderwijs en opvang van leerlingen met leesmoeilijkheden. Deze onderzoeken komen echter tot deels nogal verschillende resultaten en geven daardoor weinig houvast.

In de bestudeerde literatuur worden verschillende factoren genoemd als bevorderend of juist belemmerend voor zorgverbreding, individualisering en differentiatie in de basisschool en/of het verwijzen van leerlingen naar het speciaal onderwijs. De conclusies zijn echter niet altijd eenduidig. We vatten de uitkomsten kort samen. De conclusies ten aanzien van de *klassegrootte* zijn consistent: leerkrachten geven aan dat kleinere klassen meer mogelijkheden bieden om extra aandacht te geven aan leerlingen met problemen en diverse onderzoeken tonen ook aan dat in grotere klassen minder hulp wordt gegeven en minder wordt gedifferentieerd (Meijer, 1986; Span e.a., 1985; Reezigt e.a., 1986; Wolbert e.a., 1986). In *combinatieklassen* wordt gemiddeld meer hulp gegeven en meer gedifferentieerd dan in enkelvoudige klassen, ook al noemen leerkrachten een combinatieklas belastend en belemmerend (Span e.a., 1985; Reezigt e.a., 1986; Wolbert e.a., 1986). Veenman en Lem (1985) zijn van mening dat zorgverbreding in combinatieklassen moeilijk is te realiseren. Hun onderzoeksresultaten echter laten amper verschillen zien tussen enkelvoudige klassen en combinatieklassen, qua tijdsbesteding, didactische activiteiten en schoolprestaties (zie ook Lassche, 1986). De aanwezigheid van voldoende, geschikte *methoden en materialen* wordt door velen, waaronder leerkrachten, genoemd als voorwaarde voor het realiseren van hulp aan leerlingen met problemen en voor interne differentiatie (Ponte, 1984; Meijer e.a., 1984; Span e.a., 1985; Wolbert e.a., 1986; Reezigt e.a., 1986). Ten aanzien van *contacten tussen BAO en SO* zijn er indicaties dat deze gepaard gaan met een toename in het verwijzen (Stokking & De Vries, 1981; Span e.a., 1985; Meijer, 1986). De *visie van de leden van het schoolteam* op het leren op school, de ontwikkeling van kinderen en de samenleving is volgens Wolbert e.a. (1986) een voorname factor bij de beslissing om al dan niet te gaan differentiëren en de wijze waarop dat dan gebeurt. Diverse publicaties rapporteren positieve bijdragen van *begeleiding* (Ponte, 1984; Meijer e.a., 1984; Inspectie, 1985; Reezigt e.a., 1986). De conclusies ten aanzien van *nascholing* zijn minder éénduidig. Meijer (1986) en Reezigt e.a. (1986) zijn positief over de mogelijke bijdrage daarvan. Span e.a. (1985) stelden echter dat het gevolgd hebben van nascholing niet duidelijk samenhing met de mate waarin leerkrachten problemen zien, zich over leerlingen zorgen maken en extra hulp organiseren.

4.3 *Nascholing*

Reeds eerder (Stokking, 1988) is een overzicht gepresenteerd van literatuur over nascholing. We vatten de uitkomsten daarvan kort samen. Rond 1980 werd in Nederland voor het eerst op grotere schaal gebruik gemaakt van nascholing als beleidsinstrument van de rijksoverheid. Dit ging gepaard met een sterke voorkeur voor school- en teamgerichte nascholing. Daarvan werd een betere aansluiting verwacht bij de specifieke problemen van elke school, en bij de veranderingsprocessen die de school als sociaal systeem moet doormaken.

De eerste grootschalige cursussen in het kader van de innovatie basisschool (IVB, NIB) waren geen succes. Evaluaties spreken van het ontbreken van goede voorlichting en duidelijke werving, uiteenlopende verwachtingen bij cursisten, zeer heterogeen samengestelde cursusgroepen, sterk voorgestructureerde programma's, gebrek aan waardering voor de geboden informatie, het mislukken van cursussen door te geringe deelname en teleurstellen-

de resultaten in termen van effecten. Als belemmerende factoren noemen verschillende auteurs: te weinig tijd voor een goede voorbereiding en organisatie, te weinig rekening houden met regionale verschillen door de rijksoverheid, weinig beschikbare know-how met betrekking tot de vormgeving en inrichting van nascholing, niet berekend zijn van de opleidingsinstituten op een grootschalige aanpak, niet toegerust zijn van docenten voor de nascholingstaak, een te grote afstand tussen docenten en onderwijspraktijk, inadequate analyse van behoeften, te open of te versnipperde inhouden, ontoereikende cursusbeschrijvingen, niet geëxpliciteerde verwachtingen, problemen in de relatie PABO-OBD en motivatieproblemen op basisscholen.

Inmiddels is in Nederland meer empirisch onderzoek naar nascholing op gang gekomen. Daaruit blijkt het volgende. Van het aanbod van cursussen wordt een substantieel deel niet gegeven, vaak doordat er zich te weinig deelnemers opgeven. Als motieven of verklaringen voor het niet deelnemen wordt naast gebrek aan tijd vooral het aanbod genoemd: dit zou onvoldoende bekend zijn, onbruikbaar, te algemeen, niet passend in de activiteiten op school. Uit onderzoek blijkt verder dat de meeste cursussen niet teamgericht zijn, maar individueel; dat in ongeveer de helft van de gevallen sprake is van een intakegesprek of beginsituatievaststelling; dat er in ongeveer de helft van de gevallen geen differentiatie naar inhoud plaatsvindt; dat nascholing in de meeste gevallen bestaat uit informatieoverdracht en discussie; dat individu-gerichte en teamgerichte cursussen niet veel van elkaar verschillen; dat deelnemers aan individugerichte cursussen gemiddeld iets positiever zijn over de cursus dan deelnemers aan teamgerichte cursussen; en dat nazorg of follow-up door de nascholer of door een schoolbegeleider meestal ontbreekt. Er zijn nog weinig onderzoeken gedaan waarin kenmerken van nascholing in verband worden gebracht met resultaten, anders dan in termen van deelname en waardering. Toch worden in de literatuur veel uitspraken gedaan over wat belangrijke variabelen zijn bij nascholing resp. hoe nascholing zou moeten worden opgezet en uitgevoerd. Een nadere analyse daarvan leert dat men in de periode 1980-1987 in Nederland unaniem is in aanbevelingen, gericht op: samenhang en samenwerking tussen nascholing en begeleiding, deskundigheid van de docent, praktische gerichtheid van de cursus en steun van collega's en begeleiders bij implementatie. Onzeker is men over het belang van: de plaats van initiatief, schoolgerichte nascholing, het 'aansluiten bij behoeften', en de keuze van het aangrijpingspunt (kennis, vaardigheden, attituden, gedrag). De empirische basis voor veel uitspraken in de literatuur is echter zwak.

5 *Evaluatiecriteria*

In verschillende gesprekken met het ministerie van O&W en de LPC zijn evaluatiecriteria gespecificeerd, voorzover deze betrekking hadden op de *doelen* en *uitvoeringscondities* (zie par. 3). In Figuur 1 worden deze samengevat. Het doel 'minder verwijzingen naar het speciaal onderwijs' blijft in dit onderzoek buiten beschouwing. Daarnaar is een afzonderlijk onderzoek verricht (zie Delsing, De Jonge, Meesters, 1988). Eveneens ontbreken hier criteria m.b.t. de leesprestaties van de leerlingen. Ook de overige zijn afzonderlijk onderzocht (zie Vinjé, 1987, 1989). Deze criteria zijn door ons nader ingevuld op basis van de kennis over leesonderwijs, zorgverbreding en nascholing (zie par. 4 en par. 6.2).

Zoals bekend kan onderscheid worden gemaakt in verschillende typen evaluatiecriteria (zie ook Van de Grift, 1984; Stokking, 1984, 1989). In Figuur 2 zijn deze samengevat.

Indien we de criteria die door de opdrachtgever van het onderzoek (O&W) en de landelijke uitvoerder van het beleid (LPC) zijn gespecificeerd (Figuur 1) confronteren met het overzicht van typen criteria (Figuur 2) dan is duidelijk dat het met name gaat om proces- en produktcriteria. Deze laatste betreffen zowel het leerkrachtniveau als het schoolniveau en zowel officieel beoogde effecten (O&W) als feitelijk nagestreefde effecten (LPC). Het gaat daarbij om relatieve effecten (verbetering). (In de conclusies (par. 8) gaan we tevens kort in op de a priori criteria 'consistentie tussen gestelde doelen en gekozen middelen' en 'haalbaarheid').

Ter afsluiting van deze paragraaf merken we op dat onderscheid moet worden gemaakt tussen de keuze van evaluatiecriteria en de wijze

1 Alle of vrijwel alle PABO's blijven participeren
2 Er nemen in drie jaar tenminste 33750 leerkrachten deel
3 De motieven worden meer gebaseerd op concrete behoeften
4 De afstemming tussen vraag en aanbod wordt beter
5 De opzet en organisatie van de cursus wordt beter
6 Er worden meer competente docenten ingezet
7 De cursusinhouden worden meer praktische relevant
8 Scholen en leerkrachten nemen maatregelen t.b.v. kinderen met problemen
9 Problemen van kinderen worden in het schoolteam besproken
10 De vaardigheden van leerkrachten m.b.t. zorgverbreding worden groter
11 Men doet iets met de cursusinhoud
12 Er vindt in het schoolteam meer overleg plaats over leesproblemen
13 Er worden extra materialen aangeschaft
14 De aanpak van het leesonderwijs verandert
15 Er komt meer aandacht voor c.q. verandering in de definitie van 'problemen'
16 De ervaren problemen veranderen
17 De vaardigheden met betrekking tot leesproblemen nemen toe
18 De aanpak van leesproblemen wordt anders/beter
19 Er wordt meer gebruik gemaakt van zorgverbredende activiteiten bij leesproblemen

Figuur 1 *Evaluatiecriteria bij deze beleidsmaatregel*

waarop deze worden gemeten. Evaluatieonderzoek waarin effectcriteria worden gehanteerd (zoals hier) is 'harder' dan evaluatieonderzoek waarin bijv. alleen procescriteria worden gebruikt. Dan nog echter kan zulk effectonderzoek meer of minder 'hard' zijn, zowel qua design (zie bijv. Scheerens, 1983) als qua type meting (zie bijv. Fiske, 1971). Op deze onderzoekstechnische aspecten gaan we nu nader in.

- A priori criteria m.b.t. de te evalueren activiteit
  - consistentie tussen gestelde doelen en gekozen middelen
  - haalbaarheid
  - begrote kosten
- Procescriteria
  - keuze van uitvoerders
  - middelenkeuze (gehanteerde materialen, werkwijzen, interventies)
  - deelname (aard en omvang)
  - tevredenheid (bij uitvoerders en deelnemers)
- Produkt- of effectcriteria
  - mate waarin de doelen worden bereikt
    - door het beleid officieel beoogde
    - door de uitvoerders feitelijk nagestreefde
  - neveneffecten

  beide te onderscheiden naar:
  - (aggregatie)niveau waarop (leerling, groep, leerkracht, school, etc.)
  - effecten qua gemiddeld niveau en/of qua verdeling (spreiding)
  - absolute (vast criterium) en/of relatieve (verbetering, gezien vroeger/elders)
  - op kortere termijn en/of langere termijn (beklijving, impact)
- Doelmatigheidscriteria
  - gemaakte kosten
  - efficiency
  - verhouding kosten-baten

Figuur 2 *Overzicht van typen evaluatiecriteria*

6 *Het surveyonderzoek: opzet en uitvoering*

6.1 *Design*

Het surveyonderzoek is opgezet met een quasi-experimenteel design, met vóór-, na- en follow-up-metingen en vergelijking tussen deelnemers én niet-deelnemers. Per PABO per cursusjaar werden 15 aan de cursus deelnemende basisscholen benaderd voor beginmetingen, eindmetingen en follow-up-metingen, op basis van zelfrapportage; ook werd een controlegroep van niet-deelnemende scholen tweemaal benaderd (zie Figuur 3).

Idealiter zouden bij alle drie groepen deelnemers steeds beginmetingen èn eindmetingen èn follow-up metingen zijn gehouden. De late startdatum van het onderzoek en de vroege einddatum zorgden er voor dat bij de deelnemers '85-'86 alleen een follow-up meting kon worden gehouden en bij de deelnemers '87-'88 alleen een beginmeting en een eindmeting.

6.2 *Variabelenkeuze, operationalisering en beïnstrumentering*

De onderzoeksvariabelen zijn gekozen op grond van de doelen en uitvoeringscondities bij de nascholingsoperatie en de eerder genoemde literatuur. Dit leidde tot vier groepen variabelen:

- contextvariabelen: *algemene* kenmerken van de school en de respondent; opvattingen en activiteiten m.b.t. *nascholing*;
- beginsituatievariabelen: vaardigheden, opvattingen en activiteiten t.a.v. *leesonderwijs en zorgverbreding*;
- procesvariabelen: beschrijving en beoordeling van *de cursus en de begeleiding*; deelname, inhoud, werkwijze, ervaringen;
- effectvariabelen: *activiteiten* die volgens de respondenten hebben plaatsgevonden en/of doorgevoerde *veranderingen*; toeschrijvingen daarvan aan de cursus en/of begeleiding.

Het concept-meetinstrumentarium is voorgelegd aan een aantal deskundigen en uitgeprobeerd bij proef-respondenten. De uitkomsten daarvan zijn verwerkt in herziene instrumenten. Deze bevatten zowel losse vragen als Likert-schalen, waarvan de betrouwbaarheid steeds opnieuw is nagegaan per respondentengroep per meetmoment. In Figuur 4 zijn beide samengevat. Waar sprake is van schalen blijkt dat uit het tussen haakjes toegevoegde aantal items. Voor meer gedetailleerde informatie zie het genoemde eindrapport.

6.3 *Steekproeftrekking en respons*

Op voorhand kon worden vastgesteld dat het surveyonderzoek alleen kon worden uitgevoerd bij steekproeven, niet bij de totale populatie. Een voornaam beoordelingscriterium bij steekproeftrekking is de mate waarin belangrijke variabelen feitelijk (zullen) variëren. Indien veel variatie wordt verwacht dient de steekproeffractie groter te zijn. Variatie werd verwacht zowel tussen PABO's als tussen PABO-docenten als tussen basisscholen als tussen de afzonderlijke leden van hetzelfde schoolteam. Er waren geen gegevens bekend om de ene variatie groter te achten dan de andere.

Een tweede belangrijk criterium bij steekproeftrekking is het type generalisatie dat wordt beoogd. Indien generalisatie naar de totale populatie wordt beoogd ligt proportionele trekking het meest voor de hand (op relevante kenmerken dezelfde verhouding in de steekproef als in de populatie). Indien conclusies per subgroep nodig zijn wordt vaak eerder gekozen voor een gelijk aantal eenheden per subgroep. In dit onderzoek werd generalisatie over PABO's beoogd. Daarnaast dienden conclusies te worden getrokken per cursusjaar/scholengroep.

Een derde criterium zijn de aantallen onderzoekseenheden die minimaal nodig zijn voor het uitvoeren van statistische analyses en het verantwoord trekken van conclusies.

Een vierde criterium is gelegen in de verhouding tussen het aantal te selecteren organisaties en het aantal respondenten per organisatie. Bij een bepaald beschikbaar budget betekent méér van het één minder van het ander. Ook is relevant in hoeverre respondenten als informanten over hun organisatie kunnen worden gezien. De keuze voor bijv. alle 66 PABO's en 2 docenten per PABO zou een te versnipperd beeld opleveren. Een en ander leidde tot de beslissing om ten behoeve van het surveyonderzoek gestratificeerd-aselect 1/3 deel van de PABO's te selecteren, per PABO alle betrokken docenten te benaderen, per PABO per cursusjaar/scholengroep 1/3 deel van de betrokken scholen te selecteren en per school 1/3 deel van het team. Er zijn (voorjaar '86) 24 PABO's geselecteerd. Daarbij is gelet

Respondentengroep	1985-1986	1986-1987 sep	jan	mei	1987-1988 sep	jan	mei
Deelnemers 85-86	XXX		0				
Deelnemers 86-87		0	XXX	0		0	
Deelnemers 87-88					0	XXX	0
Niet-deelnemers		0				0	
PABO-docenten		0			0		
OBD-begeleiders		0			0		
Toelichting:	XXX	= cursus (gemiddeld 10 bijeenkomsten)					
	0	= meting					

Figuur 3 *Het design voor het survey-onderzoek*

CONTEXTVARIABELEN
SCHOOLNIVEAU
schoolgrootte
deelname OVB
leesmethode
mate zittenblijven
mate verwijzen naar SO
aanwezigheid orthotheek
LEERKRACHTNIVEAU
onderbouw/bovenbouw
klassegrootte
wel/geen combinatieklas
onderwijservaring
geslacht
ervaring met nascholing (3)
algemeen oordeel over nascholing (3)
oordeel over geschiktheid nascholing (7)
belang van de persoon van de docent (3)
belang van de inbreng van de docent (4)
belang van overzichtelijkheid aanbod (6)
belang van aandacht voor toepassing (12)

PROCESVARIABELEN
SCHOOLNIVEAU
een 7-tal motieven voor cursusdeelname
plannen voor toepassing op de school
plannen voor overdracht naar andere leerkrachten
aantal wensen bij de cursusinhoud (22)
een 8-tal beschrijvende kenmerken van de cursus (plaats, tijd, aantal, duur, omvang)
LEERKRACHTNIVEAU
een 6-tal verwachtingen over de cursus
plannen voor toepassing in de klas
mate van afdoenheid voorlichting (4)
mate waarin intake plaatsvond (8)
beoordeling persoon en inbreng docent (6)
beoordeling omvang cursusinhoud (21)
tevredenheid over de cursus (4)

BEGINSITUATIEVARIABELEN
SCHOOLNIVEAU
mate teamoverleg over leesonderwijs (5)

mate teamoverleg over leesproblemen (6)

LEERKRACHTNIVEAU
mate van gebruik van de leesmethode
belang van gestructureerd leesonderwijs (4)
mate van gebruik van extra materialen (18)
een 18-tal problemen bij leesonderwijs
een 5-tal omschrijvingen van 'leesproblemen'
het aantal kinderen met leesproblemen
een 7-tal activiteiten bij leesproblemen
oordeel over de haalbaarheid van het beleid
oordeel over juistheid keuze vak lezen
oordeel over juistheid keuze nascholing
aantal leesboeken in het lokaal

EFFECTVARIABELEN
SCHOOLNIVEAU
aanschaf van materialen en/of leesboeken en/of leesmethode en/of taalmethode
mate teamoverleg over leesonderwijs (5)
mate teamoverleg over leesproblemen (6)
LEERKRACHTNIVEAU
verandering aanpak leesonderwijs (4)
verandering aanpak leesproblemen (4)
mate gebruik van extra materialen (18)
een 18-tal problemen bij leesonderwijs
een 5-tal omschrijvingen 'leesproblemen'
het aantal kinderen met leesproblemen
een 7-tal activiteiten bij leesproblemen
oordeel over de haalbaarheid van het beleid
oordeel over juistheid keuze vak lezen
oordeel over juistheid keuze nascholing

Figuur 4 *Vragen en schalen (bij schalen tussen haakjes het aantal items)*

op de spreiding op de volgende variabelen: regionale ligging; denominatie; aantal cursusgroepen '85-'86; de samenstelling van de cursusgroepen (overwegend complete teams of individuele leerkrachten); de samenstelling van de uitvoerende docentengroep (gelet op de vak- en vormingsgebieden van de docenten); de deelname aan de LPC-voorscholing; de verschillen in cursusinhoud tussen de docenten; de aanwezigheid van werkafspraken tussen docenten en begeleiders m.b.t. de nascholing lezen; de participatie van begeleiders in de voorbereiding van de cursus. De geselecteerde PABO's mogen worden beschouwd als representatief voor alle PABO's; zie voor een gedetailleerde verantwoording het eindrapport.

De trekking van de te benaderen basisscholen verliep daarna als volgt. Elk jaar werden in de maand juni uit de van het Ministerie van O&W verkregen lijsten per PABO met de deelnemende scholen voor het daaropvolgende cursusjaar aselect 15 basisscholen getrokken bij elk van de 24 geselecteerde PABO's. Dit gebeurde via een systematische trekking met aselect begin, rekening houdend met een eventuele systematiek in de opbouw van de lijst (bij sommige PABO's waren de scholen naar gemeente en/of denominatie gerangschikt). Indien sprake was van minder dan 15 deelnemende scholen werden alle scholen geselecteerd. In de zomer van 1986 is daarnaast een ruime steekproef getrokken uit de scholen die blijkens de op dat moment beschikbare gegevens noch in 1985-1986 noch in 1986-1987 aan de cursus deelnamen. Deze scholen ontvingen in september 1986 een eerste enquête. De scholen die in 1987-1988 alsnog bleken mee te doen zijn vervolgens uit dit deelbestand verwijderd. De resterende echte niet-deelnemers ontvingen in januari 1988 een tweede enquête. Per basisschool zijn steeds drie respondenten benaderd: de directeur, een leerkracht onderbouw en een leerkracht bovenbouw.

De gemiddelde respons op alle enquêtes tesamen bedroeg 61%.

7 *Het survey-onderzoek: uitkomsten*

We beschrijven achtereenvolgens: de *beginsituatie* voorafgaande aan de cursus (7.1), de *uitvoering* van het beleid (7.2) en de gemeten *resultaten* van de nascholingsoperatie (7.3).

7.1 *De Beginsituatie*

In deze paragraaf geven we een beschrijving van de situatie met betrekking tot zorgverbreding en leesonderwijs op de basisscholen, voorafgaande aan deelname aan de cursus. We doen dit aan de hand van de resultaten van de beginmetingen bij de scholen die in '86-'87 of in '87-'88 hebben deelgenomen en de eerste meting bij de niet-deelnemende scholen. De uitkomsten bij deze drie groepen scholen waren nagenoeg identiek.

Ten aanzien van *zorgverbreding* in het algemeen vinden bijna alle leerkrachten (98%) dat kinderen met problemen extra aandacht horen te krijgen. Ook is een groot deel (87%) het er mee eens dat basisscholen verantwoordelijk zijn voor het bereiken van bepaalde minimumdoelen door alle kinderen. Tweederde (66%) twijfelt echter aan de haalbaarheid van het beleidsuitgangspunt om zoveel mogelijk alle kinderen met problemen in het basisonderwijs zelf op te vangen. Minder dan de helft (40%) is het eens met de keuze voor het vak lezen.

Ten aanzien van het *leesonderwijs* zijn bijna alle leerkrachten het er mee eens dat het hun taak is om alle kinderen te interesseren voor lezen (95%) en om alle kinderen te leren lezen (89%). Een veel kleiner aantal echter denkt dat ook alle kinderen inderdaad kunnen leren lezen (42%). We zien dat men zich gemakkelijk achter algemene idealen stelt; in de praktijk echter zijn velen sceptisch over de realiseringsmogelijkheden. De *ervaren problemen bij leesonderwijs* zijn de volgende. De helft van de leerkrachten (53%) weet wel welke kinderen welke problemen hebben en wat zij daaraan kunnen doen, maar heeft daar onvoldoende tijd voor. Een kwart van de leerkrachten weet niet om welke problemen het gaat (23%) of wat zij er aan kunnen doen (22%). De helft heeft behoefte aan meer praktische tips voor de leeslessen, een kwart aan observatie- en toetsmaterialen en geschikte lees-/oefenstof (qua moeilijkheidsgraad en aantrekkelijkheid). Problemen met de eigen (theoretische) kennis, de organisatie van leeslessen of de beschikbare leesmethode worden weinig vermeld.

De meeste leerkrachten spreken van *leesproblemen* als een kind duidelijk achterblijft bij zijn klasgenoten (71%) of bij wat het,

gelet op zijn mogelijkheden, zou moeten kunnen (63%). In mindere mate spreken leerkrachten van leesproblemen als het lezen bij andere vakken problemen oplevert (48%), als een kind geen interesse heeft in lezen (27%) of als zij teveel aandacht moeten besteden aan een bepaald kind (19%). Het wordt amper als probleem gezien als een groot deel van de klas duidelijk achterblijft bij het tempo dat in de methode staat aangegeven. De *afspraken in het team* betreffen vooral het afnemen van toetsen (66%) en het houden van leerlingbesprekingen (55%). Afspraken over andere zaken, zoals de aansluiting tussen voorbereidend, aanvankelijk en voortgezet lezen, observatie, foutenanalyse en extra instructie, komen minder voor (25-40%). De *situatie in de klas* is als volgt. De leerkrachten hebben gemiddeld 3 kinderen in de klas die naar hun zeggen problemen met lezen hebben of waarschijnlijk zullen krijgen. Driekwart van de leerkrachten geeft deze kinderen extra aandacht tijdens de leeslessen zelf, terwijl de rest van de klas doorwerkt. De helft van de leerkrachten geeft hen aanvullende oefenstof. Extra aandacht na schooltijd, apart toetsen en hulp van een remedial teacher worden elk door eenderde deel van de leerkrachten gemeld. Toetsen van de vorderingen van de klas als geheel doet eenderde deel slechts eens per halfjaar of minder.

7.2 *De uitvoering van het beleid*

7.2.1 De landelijke ondersteuning

De LPC ontwikkelden 18 materialen ten behoeve van cursus A. De materiaalontwikkeling vond plaats bij of namens de drie landelijke centra: het APS, het CPS en het KPC. Deze materialen werden in de periode januari-mei 1986 op zogenaamde 'voorscholingsdagen' gepresenteerd aan belangstellende PABO-docenten en OBD-begeleiders. Enkele honderden personen van de meeste PABO's en OBD's in ons land, namen aan deze dagen deel. De materialen werden in aparte sessies toegelicht. De belangstelling daarvoor varieerde van 7% tot 76%. Uit de surveys bij PABO's en OBD's bleek dat tweederde van de docenten en de helft van de begeleiders de inhoud van de voorscholing enigszins interessant vond. Tweederde van de docenten en eenderde van de begeleiders vond de inhoud enigszins relevant. Docenten misten vooral suggesties over hoe de cursus inhoudelijk, organisatorisch en didactisch zou kunnen worden aangepakt en ook had men meer specifieke cursusmaterialen willen hebben. Begeleiders misten vooral een visie op en een kader voor deze nascholingscursus. De meeste materialen bleken bekend bij en in bezit van 60-80% van de docenten en 30-50% van de begeleiders. Er werd echter weinig of geen gebruik van gemaakt. Slechts een vijftal van de door de LPC uitgebrachte materialen werd door tenminste 40% van de nascholers en/of tenminste 40% van de begeleiders redelijk veel of veel gebruikt. Deze hadden vooral betrekking op differentiatie en individualisering. Daarnaast werden met name remediërende materialen van onder meer educatieve uitgevers redelijk veel of veel gebruikt.

7.2.2 De organisatie en inhoud van de cursussen

De in het onderzoek betrokken PABO's verzorgden de cursus in de drie opeenvolgende cursusjaren voor gemiddeld 12 resp. 15 resp. 6 cursusgroepen (het aantal varieerde van 1-30). Daarvoor werden gemiddeld 5 docenten ingezet. De cursusgroepen bestonden elk jaar uit gemiddeld 21 cursisten. De groepen waren meestal (75%) samengesteld uit complete schoolteams en voor de rest uit individuele leerkrachten of teams en individuele leerkrachten. De meeste cursussen werden na schooltijd gegeven, op één van de deelnemende scholen of op de PABO zelf. Gemiddeld besloeg een cursus 10 bijeenkomsten (variatie: 4-18) van 3 uur (variatie: 1-4). Er vond meestal vooraf overleg plaats met de cursisten. Dit betrof het eerste jaar vooral plaats en tijdstip van de cursus, in latere jaren ook meer de inhoud.

Ten aanzien van deze inhoud bleek dat de docenten een groot aantal onderwerpen behandelden. Het meest genoemd werden de theorie over en de problemen die kunnen optreden bij het leren lezen. Er was relatief weinig aandacht voor de aanpak van het leesonderwijs op de school als geheel, evenals voor specifieke zorgverbredingsactiviteiten zoals leerlingbesprekingen. Een groot maar over de jaren afnemend aantal docenten zei met de cursus te beogen dat de vaardigheden van leerkrachten in het hanteren van leesproblemen toenemen (60-45%), een toenemend aantal te beogen dat het inzicht van leerkrachten toe-

neemt (35-45%). De deelnemende leerkrachten vonden echter blijkens het onderzoek veel minder dat de cursus gericht was op vaardigheidsvergroting (25%) en juist in meerderheid dat deze gericht was op kennisvergroting (70%).

7.2.3 De deelname aan de cursussen

In de cursusjaren '85-'86 en '86-'87 hebben alle (aanvankelijk ruim 60) PABO's cursussen A verzorgd. In het cursusjaar '87-'88 was er nog sprake van 41 PABO's; daarvan verzorgden er 33 in dat jaar cursus A. Daarnaast verzorgden bijna alle PABO's met ingang van 1-1-'87 ook cursussen B (begrijpend lezen) en een 15-tal PABO's met ingang van 1-1-'88 ook cursussen C (taalvaardigheden). De omvang van de aanmelding vanuit de basisscholen was als volgt (uitgedrukt in personen, afgerond op 1000-tallen): voor cursus A in het eerste jaar 16000, in het tweede jaar eveneens 16000, in het derde jaar 5000. Voor cursus B in het tweede jaar 4000 en in het derde jaar 5000. Voor cursus C in het derde jaar 1000.

Uit het onderzoek blijkt dat ongeveer 20% van de scholen die zich in het voorjaar (in één van de drie jaren) aangemeld hadden bij een PABO niet daadwerkelijk in het daarop volgende cursusjaar aan de cursus ging deelnemen. Elk cursusjaar gaf daarnaast 13% van de respondenten aan dat op hun school voortijdig één of meer leerkrachten met het volgen van de cursus waren gestopt. Daarnaast gaf in de eerste twee cursusjaren 1% van de respondenten aan dat dit het geval was met alle deelnemers van de school. Deelname aan cursus A in een bepaald jaar wil niet zeggen dat niet ook sprake kan zijn van deelname aan cursus A in een eerder of later jaar. Het totaal aantal *verschillende* personen dat in deze drie jaren deelnam aan cursus A was ongeveer 25000. Deelname gebeurde meestal met het hele schoolteam (80%).

In het eerste jaar was het dominerende motief om deel te nemen dat men iets verwachtte van de cursus (75%). Genoemd motief speelde in het tweede en derde jaar een minder grote rol (57% resp. 47%). Steeds ongeveer de helft van alle respondenten noemde de extra formatie als deelnamemotief. *Behoefte* aan een dergelijke cursus werd constant gemeld door slechts 15% van de deelnemers. Bij de scholen die in de onderzochte periode van drie jaar *niet* aan de cursus deelnamen waren de motieven daarvoor met name, en in toenemende mate, gebaseerd op negatieve berichten over de cursus en op een relatieve voorkeur voor samenwerking met de OBD (aan een invoeringsprogramma).

De extra formatie is in ongeveer 10% van de gevallen gebruikt voor het volgen van de cursus. In ongeveer 65% van de gevallen is de extra formatie zonder meer toegevoegd aan de formatie van de school. Op een beperkt aantal scholen is de extra formatie gebruikt voor hulp aan individuele leerlingen en/of voor het verrichten van andere specifieke taken.

Ongeveer 65% van alle scholen ontving tijdens één of meer van deze drie jaren begeleiding bij lezen. Ongeveer 20% van alle leerkrachten ontving tijdens één of meer van deze drie jaren individuele begeleiding bij het leesonderwijs.

7.2.4 De waardering voor de cursussen

De meeste docenten (80-90%) waren positief over de inhoud, werkwijze en organisatie van door hen gegeven cursussen. Van de deelnemende leerkrachten was in het eerste jaar slechts een minderheid positief over deze aspecten (45%). De waardering nam echter toe: in het derde jaar was over elk aspect 65% van de deelnemers positief. Over de rol van de schoolbegeleider tijdens de cursus was slechts een minderheid van de docenten tevreden (35%). Onder de leerkrachten was ook hierover een toenemend deel tevreden (van 45% naar 65%).

7.3 *De gemeten resultaten van de naschollingsoperatie*

7.3.1 Resultaten volgens de deelnemers aan de cursus

De helft van de leerkrachten heeft *iets gedaan met de cursusinhoud* in de eigen klas. Genoemd worden (elk door 15-35%): aandacht aan voorlezen en stillezen, de inrichting van een kijk-luister- of lees-schrijfhoek, aandacht besteden aan leesproblemen, gebruik maken van extra oefenstof en remediërend materiaal, maken van vorderingenoverzichten en foutenanalyses. De andere helft heeft er niets mee gedaan in de eigen klas, omdat de cursusinhoud niet van toepassing was (doordat men zelf geen aanvankelijk leesonderwijs verzorgde) en/of omdat men de cursusinhoud negatief beoordeelde.

Een deel van de leerkrachten stelt daarnaast dat er *iets is veranderd*. Ongeveer 40% van de deelnemers stelt dat men meer inzicht heeft gekregen in het leesproces en in leesproblemen, ongeveer 15% dat men er anders over is gaan denken en ruim 10% dat men vaardiger is geworden in het hanteren van leesproblemen. Ten aanzien van concrete activiteiten kan het volgende worden gemeld. Veranderingen bij het geven van extra aandacht aan kinderen met problemen, het afnemen van toetsen en het geven van extra oefenstof worden elk door 35% van de leerkrachten gemeld. Veranderingen in de manier waarop de leesmethode wordt gebruikt, de tijd die aan diverse leesvormen wordt besteed, de mate van differentiatie en het inschakelen van extra hulp worden na afloop van de cursus door 10% van de deelnemers genoemd, in het daaropvolgende schooljaar door 20%. Dergelijke veranderingen vergen kennelijk meer tijd.

Tevens zijn er *materialen aangeschaft*. Ook hiervan was vaak pas sprake in het op de cursus volgende schooljaar. Daarbij ging het 't vaakst om de aanschaf van nieuwe leesboeken (60%) en extra (remediërende) materialen (50%). In mindere mate betrof dit de aanschaf van een nieuwe taal- of leesmethode (20% resp. 10%). Dit laatste werd slechts door een minderheid van de leerkrachten *toegeschreven aan de cursus of de begeleiding*. De aanschaf van extra materialen werd echter door ongeveer 70% toegeschreven (mede) aan de cursus, door ongeveer 20% (mede) aan de begeleiding en door ongeveer 20% (mede) aan ander factoren. Voor de bovengenoemde veranderingen in concrete activiteiten gelden dezelfde percentages. Hierbij moet worden bedacht dat slechts een beperkt deel van de leerkrachten begeleiding ontving bij het leesonderwijs (zie 7.2.3).

7.3.2 Resultaten die kunnen worden afgeleid uit vergelijkingen tussen deelnemende en niet-deelnemende scholen en verschuivingen op deelnemende scholen in de loop van de tijd De uitkomsten in de verschillende scholengroepen (deelnemers 1e, 2e en 3e jaar en niet-deelnemers) en in de verschillende uitgevoerde meetrondes komen over de hele linie in sterke mate overeen. Vergelijkend tussen de metingen (begin-, eind-, en follow-up-metingen) kan een aantal verschuivingen worden geconstateerd, waarbij we als criterium nemen een verschuiving van tenminste 10% in afzonderlijke groepen of een verschuiving van 5% die optreedt in alle groepen in dezelfde richting.

In de *meningen* van leerkrachten nemen bij de niet-deelnemers de twijfels aan de haalbaarheid van het beleid af (van 67% naar 56%). Bij de deelnemers neemt de mate waarin de keuze voor een speerpunt lezen juist wordt geacht toe (van 40% tot 55%). De overige verschuivingen in meningen doen zich zowel bij de deelnemers als niet-deelnemers voor, te weten afname van de mate waarin men het tot de eigen taak rekent om alle kinderen te leren lezen (!), meer praktische tips wil hebben voor leeslessen, te weinig observatie- en toetsmateriaal heeft en het moeilijk vindt om geschikte lees/oefenstof te vinden.

In de *activiteiten* van de leerkrachten constateren we een afname van de mate waarin men bij leesproblemen de hele klas pas op de plaats laat maken en van de inschakeling van een remedial teacher, en een toename van het gebruik van extra materialen en van het geven van aanvullende oefenstof. Bij de hulp bij het leesonderwijs zien we een afname van de mate waarin men hulp krijgt van een remedial teacher en een toename van de hulp van een begeleider.

De meeste genoemde verschuivingen waren niet veel groter dan het gestelde minimumcriterium van ongeveer 10% en traden op bij zowel de deelnemers als de niet-deelnemers. De hier niet genoemde meningen en activiteiten lieten geen verschuiving zien. Zo trad er bijvoorbeeld geen verschuiving op in de mate waarin leesonderwijs en zorgverbreding onderwerp waren van overleg of besluitvorming in het schoolteam.

Er zijn variantieanalyses uitgevoerd om te toetsen op verschillen tussen groepen scholen en om na te gaan of er ten aanzien van leesonderwijs en zorgverbreding op deelnemende scholen tussen begin- en eindmeting méér is veranderd dan op niet-deelnemende scholen. Dit was niet het geval. Er zijn ook nog multiple regressieanalyses uitgevoerd om na te gaan in hoeverre procesvariabelen te voorspellen zijn vanuit context- en situatiekenmerken en in hoeverre effectvariabelen te voorspellen zijn vanuit context- en situatiekenmerken en procesvariabelen samen. De gevonden verbanden waren zwak en bovendien niet constant over de onderscheiden scholengroe-

pen. De gemeten effecten zijn dus niet alleen klein maar ook niet goed te herleiden tot bepaalde bevorderende of belemmerende factoren.

8 *Conclusies*

De hier te trekken conclusies betreffen de beleidsmaatregel op zich, de gemeten effecten van de beleidsmaatregel en het uitgevoerde surveyonderzoek.

In hoeverre is het beleid (par. 2) afgestemd op de kennis over leesonderwijs, zorgverbreding en nascholing (par. 4)? De literatuur over leesonderwijs mondt uit in de conclusie dat variatie en differentiatie nodig zijn, en daarmee een ruime keus uit materialen, en deskundige en vaardige leerkrachten. Aangezien SLO, CITO, Universiteiten en Educatieve uitgevers regelmatig nieuwe materialen ontwikkelen en uitbrengen lijkt het accent in het beleid op bekwaamheidsvergroting van leerkrachten terecht. De literatuur over zorgverbreding leidt eveneens tot de conclusie dat diagnostiserende instrumenten, remediërende materialen en de inzet van ervaren leerkrachten nodig zijn. Daarnaast zijn maatregelen op schoolniveau nodig. De uitwerking daarvan in groepen van 20 cursisten waarin verschillende schoolteams vertegenwoordigd zijn ligt minder voor de hand. Ten aanzien van de keuze voor nascholing ten slotte kan men zich afvragen of de operatie uitvoerbaar kon worden geacht, gezien de bekende structurele problemen (o.a. in de relatie PABO's-OBD's) en de relatief beperkte kennis van PABO-docenten ten aanzien van problemen bij het leren lezen. Wat waren de gemeten effecten van de beleidsmaatregel? De dominerende uitkomst van het surveyonderzoek is de grote overeenkomst op veel variabelen tussen alle metingen, ongeacht groep of tijdstip. Dit betekent enerzijds dat er tussen vroege deelnemers en late deelnemers en niet-deelnemers geen grote verschillen zijn, anderzijds dat er weinig sprake is van verandering, of men nu wel of niet de cursus heeft gevolgd.

In Figuur 1 zijn de criteria genoemd die volgens O&W en de LPC gehanteerd zouden moeten worden bij de evaluatie van de cursus. Twee criteria zijn blijkens de resultaten van dit onderzoek duidelijk gehaald: de PABO's zijn blijven participeren in de nascholing en de afstemming tussen vraag en aanbod werd geleidelijk beter. Ook ten aanzien van een derde criterium kan men de balans laten doorslaan naar de positieve kant: ook de opzet en organisatie van de cursus leken geleidelijk te worden verbeterd. Over een vierde criterium, namelijk of er meer competente docenten worden ingezet, kan op grond van de door ons verzamelde gegevens niets worden gezegd.

Ten aanzien van een aantal criteria, namelijk die betreffende maatregelen en activiteiten op school- en klasniveau met betrekking tot zorgverbreding en betreffende het gebruik maken van de cursusinhoud, de aanschaf van extra materialen, het vaardiger worden van leerkrachten en het veranderen van de aanpak, is geen duidelijke uitspraak mogelijk. Dergelijke resultaten werden weliswaar door gemiddeld 35% van de respondenten zelf als effecten van de cursus genoemd; wij constateren echter slechts verschuivingen van ongeveer 10%, welke in veel gevallen óók optreden op de scholen die niet aan de cursus deelnamen en dus eerder beschouwd moeten worden als een algemeen speerpunteffect dan als een specifiek cursuseffect.

Enkele criteria werden, gezien de uitkomsten van dit onderzoek, niet gehaald: het aantal deelnemers aan de cursus bleef duidelijk achter bij het gestelde criterium, de motieven van degenen die wèl deelnamen werden niet van jaar tot jaar meer gebaseerd op concrete behoeften, de cursusinhouden werden niet van jaar tot jaar duidelijk meer praktisch relevant (ten behoeve van zorgverbreding), en het overleg in het schoolteam over leesproblemen nam gemiddeld niet toe.

Wat is tenslotte de waarde van het uitgevoerde surveyonderzoek? Er is niet alleen gekeken naar procesgegevens maar ook naar evaluatiecriteria m.b.t. effecten. Zoals eerder gezegd kan ook effectonderzoek nog meer of minder "hard" worden uitgevoerd. De in dit onderzoek gehanteerde quasi-experimentele opzet en het gebruik van zelfrapportage leiden niet tot de meeste harde conclusies die men zou kunnen wensen. Toch menen wij dat er voldoende betrouwbare en valide evidentie is verkregen voor de hierboven getrokken conclusies over de effectiviteit van de onderzochte beleidsmaatregel.

Literatuur

Alm, R.S. The educational causes of reading difficulties. *Journal of research and development in education*, 1984, *14*, 4, 41-49.

ARBO, *Het moet ons een zorg zijn. Advies over zorgverbreding in het basisonderwijs*. Zeist: mei 1984.

ARBO, *Aanvullende adviserende notitie over zorgverbreding in het basisonderwijs*. Zeist: september 1984.

ARBO, *Advies m.b.t. Raamplan nascholing zorgverbreding. Thema 2: Begrijpend lezen*. Zeist: februari 1986.

Carter, S.M., Materials and the disabled reader in the elementary school. *Journal of research and development in education*, 1984, *14*, 4, 8-84.

Chauveau, G., E. Rogovas-Chauveau, Les processus d'acquisition ou d'échec en lecture au cours préparatoire. *Revue Francaise de Pédagogie*, 1985, nr. 70, janvier/fevrier/mars, 5-10.

Cronbach, L.J., *Designing of evaluations of educational and social programs*. San Fransisco: Jossey-Bass, 1982.

Delsing, P.J., J.F.M. de Jonge, M. Meesters, *Leerlingverwijzingen van basisonderwijs naar speciaal onderwijs*. Leiden: Research voor Beleid, 1988.

Dongen, D. van, *Leesmoeilijkheden. Naar diagnostiserend onderwijzen bij het leren lezen*. Tilburg: Zwijsen, 1984.

Dongen, D. van, 'Lezen is voor 99% een individuele aangelegenheid'. Interview. *School en begeleiding*, 1985, 11, 9-12.

Fiske, D.W., Measuring the concepts of personality. Chicago: Aldine Atherton, 1971.

Grift, W. van de, Problemen bij summatief evaluatie onderzoek. *Pedagogische Studiën*, 1984, *61*, 3, 104-115.

Groot, A.D. de, *Begrip van evalueren*. Den Haag: VUGA, 1986.

Glass, G.V. & F.S. Ellet, Evaluation research. *Annual Review of Psychology*, 1980, *31*, 211-228.

Hofstee, W.K.B., *De empirische discussie*. Meppel: Boom, 1980.

House, E.R., *The logic of evaluative argument*. Los Angeles: CSE Monograph Series in Evaluation, 1977.

Inspectiewerkgroep "schoolbegeleiding en zorgverbreding", *Schoolbegeleiding en Zorgverbreding. Een verkennend onderzoek naar de vertaling van het zorgbreedtebeleid door de onderwijsbegeleidingsdiensten. Stand van zaken kalenderjaar 1985*. Rijswijk: Inspectie, 1985.

Jonge, J. de, M.P. Vinjé, K.M. Stokking, B. Span, *De nascholingsoperatie 'speerpunt lezen'. Zorgverbredingsbeleid geëvalueerd*. Leiden: Research voor Beleid, 1989.

Laarhoven, A. van, Moeilijk lezende kinderen geen probleem op zich! *School en Begeleiding*, 1985, 7, 22-24.

Lassche, K.J., *Combinatieklassen en ervaringen van leerlingen*. Nijmegen: Vakgroep onderwijskunde, 1986.

Meijer, C.J.W., *De rol van diagnostiek van leermoeilijkheden in onderwijsbegeleiding. Interimrapport deel 3: verslag van vijf gesprekken met leerlingbegeleiders*. Groningen: Vakgroep onderwijskunde, 1982.

Meijer, C.J.W., *Verwijzing en terugplaatsing. Verslag van resultaten van het onderzoek naar verwijzing en terugplaatsing van leerlingen van basisonderwijs naar speciaal onderwijs en vice versa.s, Leerlingverwijzingen van basisonderwijs naar speciaal onderwijs*. Leiden: Research voor Beleid, 1988.

Meijer, C.J.W., A.J. Verwilligen-Prummel, Chr. Rippen, *Projecten verwijzing en toelating en projecten terugplaatsing*. Groningen: RION, 1984.

Melis, G., Minimum-doelstelling voor het lezen in de basisschool. *School en Begeleiding*, 1985, 7, 24-29.

Ministerie van O&W, *Concept-beleidsnotitie zorgverbreding*. Den Haag: 1984a.

Ministerie van O&W, *Beleidsnotitie zorgverbreding*. Den Haag: 1984b.

Ministerie van O&W, *Nascholing Zorgverbreding. OCA-nascholingscursus speerpunt leesonderwijs*. Circulaire BO/VA. Den Haag: 1985a.

Ministerie van O&W, *Werken aan de basisschool, 1985-1986 en 1986-1987. Overzicht van maatregelen en aktiviteiten*. Den Haag: 1985b.

Ministerie van O&W, *Verslag van overleg O&W-PABO's februari 1986*. Zoetermeer: 1986a.

Ministerie van O&W, *Verslag van een onderzoek naar het verloop van de nascholingscursus "Speerpunt lezen". Rapport van de inspectie*. Zoetermeer: 1986b.

Mommers, M.J.C., Een taakanalytische benadering van het leren lezen en diagnostiserend onderwijzen. *Tijdschrift voor Taalbeheersing*, 1985a, *7*, 4, 281-290.

Mommers, M.J.C., Zorgverbreding, speerpunt lezen. *School en Begeleiding*, 1985b, 7, 4-8.

Mommers, M.J.C., C.A.J. Aarnoutse, Achtergronden voor de constructie van leesmethoden voor het basisonderwijs. *Pedagogisch Tijdschrift*, 1984, *9*, 3, 139-148.

Mommers, M.J.C., C.A.J. Aarnoutse, B.W.G.M. Smits, Het kunnen structureren van visuele en auditieve gegevens in relatie tot leren lezen en spellen. *Pedagogische Studiën*, 1985, *62*, 280-296.

Mommers, M.J.C., D. van Dongen, Het voorspellen van lees- en spellingprestaties in het eerste leerjaar. *Pedagogische Studiën*, 1984, *61*, 153-164.

Mommers, M.J.C., J.F.J. van Leeuwe, J.H.L. Oud, J.M.A.M. Janssens, *Een longitudinaal onderzoek naar de samenhang tussen technisch lezen, begrijpend lezen en spelling in de eerste drie leerjaren van de lagere school*. Nijmegen: Vakgroep onderwijskunde, 1985.

Pink, W.T., R.E. Leibert, Reading instruction in the elementary school: a proposal for reform. *The elementary school journal*, 1986, *87* 1, 51-67.

Ponte, P., Zorgbreedte en onderwijsbegeleiding. *VOU-blad*, 1984, 2-10. Utrecht: Vakgroep onderwijskunde, 1984.

Reezigt, G.J., M.H. van Dijk, J.J.F. Bosveld, *Differentiatie op de basisschool*. Den Haag: SVO, 1986.

Roelofs, E.C., J.J.C.M. Visser, T. Boland, M.J.C. Mommers, *Leesvaardigheidscomponenten in de basisschool: hoe is hun onderlinge relatie na zes jaar leesonderwijs?* Paper voor de ORD 1987. Groningen: 1987.

Scheerens, J., *Evaluatie-onderzoek en beleid*. Den Haag: SVO, 1983.

SLO, *Leren lezen en schrijven in het schoolwerkplan*. Enschede: 1981.

SLO, *Leesonderwijs in kaart gebracht*. Enschede: 1986.

Smits, B., M. Mommers, C. Aarnoutse, Geslachts- en leeftijdsverschillen bij het leren lezen en spellen. Een longitudinale studie. *Pedagogische Studiën*, 1985, *62*, 415-427.

Span, B., I.M. Abbring, C.J.W. Meijer, *De opvang van leerlingen met problemen in het regulier onderwijs*. Den Haag: SVO, 1985.

Stokking, K.M., *Toetsend onderzoek*. Dissertatie. Groningen: 1979.

Stokking, K.M., Evaluatieonderzoek. In: P.G. Swanborn & L. Rademaker (red.), *Sociologische grondbegrippen*, Utrecht: het Spectrum, 1982, 219-236.

Stokking, K.M., *Interpreteren en evalueren*. Deventer: Van Loghum Slaterus, 1984.

Stokking, K.M., *Hoe bereiken we de scholen?* Tilburg: Zwijsen, 1986.

Stokking, K.M., Nascholing basisonderwijs: onderzoek 1980-1987. *Pedagogische Studiën* 1988, *65*, 9, 363-367.

Stokking, K.M., *Evaluatiecriteria bij onderwijsbegeleiding*. Paper gepresenteerd tijdens de Universitaire Zomercursus Onderwijsbegeleiding. Zeist: 1989.

Stokking, K.M. & F.J. Leenders, *Effectiviteit van nascholing en nascholingsbeleid. Eindrapport van het evaluatieonderzoek naar de landelijke nascholingsoperatie speerpunt lezen*. Den Haag: SVO, 1988.

Stokking, K.M., A.K. de Vries, *Een huis in de pels. Eindverslag van het GEON-project*. Den Haag: Staatsuitgeverij, 1981.

Stufflebeam, D.L. (Ed.), *Educational evaluation and decision making*. Itasca: Peacock, 1972.

Theunissen, J., *Project nascholing zorgverbreding lezen basisonderwijs*. Eerste evaluatierapport voorscholingsdagen. KPC, Den Bosch: 1986a. Idem, tweede rapport, 1986b.

Veenman, S., P. Lem, Combinatieklas vormt zware belasting voor leerkracht. Geen tijd voor zorgverbreding of differentiatie. *Didaktief*, 1985 (december) 13-15.

Vernooy, K., Aandachtspunten bij de begeleiding van zorgverbreding lezen in de school. *School en Begeleiding*, 1985, 7, 13-17.

Vinjé, M.P., *Opzet en voormeting van het landelijk onderzoek leesvaardigheid*. Arnhem: CITO, 1987.

Vinjé, M.P., *Ontwikkelingen in de leesvaardigheid van leerlingen in het basisonderwijs. Resultaten van het Landelijk Onderzoek Leesvaardigheid in 1986 en 1988*. Arnhem: CITO, 1989.

Vinke, J., P. Reitsma, A. van der Leij, Subtypen van leeszwakke kinderen? Een clusteranalytisch onderzoek. *Pedagogische Studiën*, 1985, *62*, 329-338.

Vos, A., Zorgverbreding, doen! *Jeugd in School en Wereld* 1986, *70*, 8, 36-39.

Werts, M., Omgaan met verschillen in het leesonderwijs. *School en Begeleiding*, 1985, 7, 29-32.

Weterings, A., C. Aarnoutse, *Onderwijs in begrijpend lezen. Een onderzoek naar leerkrachtactiviteiten in het vierde leerjaar*. Paper voor de ORD 1986. Utrecht: 1986.

Winnubst, T., Hoe voortgezet is het voortgezet leesonderwijs? *School en Begeleiding*, 1985, 7, 18-22.

Wolbert, R.G.M., I. Schaap-Hummel, P. Span, *Individualisering en differentiatie in de basisschool*. Utrecht: Vakgroep onderwijskunde, 1986.

Curricula vitae

K.M. Stokking is senioronderzoeker bij de Vakgroep Onderwijskunde van de Rijksuniversiteit Utrecht en het ISOR, Afdeling Onderwijsonderzoek. Hij leidt het onderzoeksprogramma 'Onderwijsondersteuning' en de gelijknamige onderzoeksgroep. Zijn onderzoeksactiviteiten betreffen verspreiding en gebruik van informatie en produkten in het onderwijs, de aansluiting tussen vraag en aanbod, en evaluatieonderzoek van beleidsmaatregelen en opleidingsprogramma's.

F.J. Leenders studeerde psychologie aan de Rijksuniversiteit Utrecht met als afstudeerrichting organisatiepsychologie. Zij werkt sinds 1980 bij de Vakgroep Onderwijskunde van de Rijksuniversiteit Utrecht en het ISOR, Afdeling Onderwijs-

onderzoek. Zij is lid van de onderzoeksgroep *Onderwijsondersteuning*. Haar onderzoek betreft evaluatieonderzoek van beleidsmaatregelen en onderzoek naar verspreiding en gebruik van informatie en produkten.

Adres van de auteurs: Afdeling Onderwijsonderzoek ISOR / Vakgroep Onderwijskunde, Rijksuniversiteit Utrecht, Postbus 80.140, 3508 TC Utrecht.

Manuscript aanvaard 5-1-'90

Summary

Stokking, K. M., F. J. Leenders, 'Implementation and effects of the Dutch national program "inservice training reading instruction".' *Pedagogische Studiën*, 1990, *67*, 245-260.

In 1984 the Dutch Ministry of Education decided to start a national program on training teachers of primary schools to help children who fail to achieve the learning goals and have persistent learning-problems. During several years every primary school could follow an in-service training course at a local Pedagogical Acadamy. During the first three years the courses were focussed on reading-difficulties en reading-instruction. At the same time the Ministry gave order to evaluative research of the program, in order to get information for possible corrections. From 1986 to 1988 an evaluative research study was made of the courses and the results of the program. The main conclusion is that the effects of the courses on opinions, on activities in the classrooms and on activities at school-level, were minimal. There were hardly any differences between teachers who participated in the courses and teachers who did not.

De wereld in ons hoofd

Een onderwijsexperiment in verband met het ruimtelijke voorstellings- en structureringsvermogen

A. VERHETSEL
Universitaire Faculteiten Sint Ignatius, Antwerpen

Samenvatting

In vele alledaagse situaties worden we met ruimtelijke problemen geconfronteerd, zoals wanneer we de weg zoeken naar een nieuwe winkel of de nieuwe woonplaats van een vriend of wanneer we omwille van files in de stad een alternatieve route uitproberen. We leven bovendien in een tijd met snelle en overdadige informatiestromen uit alle hoeken van de wereld. Zelden slaan we deze ruimtelijk informatie bewust op in het geheugen. De geografische kennis van velen is dan ook erg verwarrend. Hier rapporteren we het ontwerp, de organisatie en de besluiten van een onderwijsexperiment in verband met ruimtelijk redeneren. Wanneer twaalfjarige kinderen inzicht verwerven in de ruimtelijke basisconcepten, nl. perspectief, schaal, oriëntatie en symboolvoorstelling, kunnen ze mentaal ruimtelijke voorstellingen construeren (cognitieve kaarten) die hen toelaten goede oplossingen te vinden voor ruimtelijke problemen op geografische schaal.

1 *Inleiding*

De ruimtelijke dimensie is aanwezig in het alledaagse. We verplaatsen ons, we gaan op vakantie, we hebben familie in andere streken wonen. We zijn gefascineerd door satellieten en de 'space shuttle'. We luisteren naar musici uit alle delen van de planeet, auto's en computers uit Azië omringen ons. We moeten bewust zijn van de ruimtelijke dimensie in informatie en ons uitrusten om deze, vaak verwarde informatie aan te kunnen. Dit is ongetwijfeld een taak voor het aardrijkskundeonderwijs. Maar vooraleer uit de geografische invalshoek aan sociaal-politieke vorming, attitudeontwikkeling en ontplooiïng kan worden gedaan, moet er inzicht in en affiniteit zijn voor de 'ruimtelijke dimensie'. Een goede mentale ruimtelijke voorstelling kan als ophangbord fungeren voor zinvolle feitenkennis en zo de start betekenen voor ruimtelijke bewustwording. Leerlingen bekommeren zich niet spontaan om de ruimtelijke dimensie van een probleem. Het belang van de plaats van voorwerpen te kennen ziet men niet zo maar in. Systematisch onderwijs is noodzakelijk in plaats van een toevallig aanleren. Uit onderzoek (Verhetsel, 1989) blijkt dat ruimtelijke kennis slechts zeer incidenteel wordt verworven en dat de meeste mensen niet in staat zijn een goede initiële ruimtelijke probleemrepresentatie op te bouwen. In deze publikatie melden we de opbouw, het verloop en de besluiten van een experimentele lessenreeks ter verbetering van de ruimtelijke voorstellings- en structureringsvaardigheid.

2 *Opbouw van de experimentele lessen*

2.1 *Perspectief, oriëntatie, schaal en symbolisatie: basisconcepten voor de opbouw van een efficiënte ruimtelijke voorstelling*

Uit de studie van cognitief geografische en onderwijsgeografische publikaties blijkt dat de resultaten van onderzoek naar een efficiënte cognitieve geografische voorstelling zeer kort samen te vatten zijn: hoe meer ervaring iemand heeft met geografische problemen, hoe beter zijn oplossingsvermogen voor geografische problemen. De variabelen die men onderzoekt (leeftijd, periode dat men een omgeving kent, het meest gebruikte vervoermiddel, de complexiteit van de fysische omgeving), zijn immers alle te herleiden tot min of meer actieve ervaring met de omgeving hebben. Een meer fundamentele vraag is: hoe wordt de verworven ervaring opgeslagen, verwerkt en gebruikt?

Leerpsychologen leggen de nadruk op een goede probleemrepresentatie om de vakin-

Tabel 1 *De ruimtelijke begrippen aan de basis van het kaartbegrip: abstractieniveaus*

	PERSPECTIEF	SCHAAL	ORIENTATIE	SYMBOLISATIE	KAARTBEGRIP
MATERIEEL NIVEAU	wereld ziet er anders uit naargelang standpunt	op terrein schaalverkleining	situeren t.o.v. vaste herkenningspunten	gelijkaardige zichtbare verschijnselen groeperen	aandacht voor de plaats van dingen
CARTOGRAFISCH NIVEAU	verticaal perspectief	verkleinen volgens vaste verhouding	kompasrichtingen	werken met legende	lokalisatie zoeken en registreren
MENTAAL NIVEAU	landschap op kaart visualiseren naar eigen gezichtspunt en landschap vanuit verticaal oogpunt kunnen voorstellen	inzicht hebben in verkleining van kaarten en gepaste, constante verkleining voor kaart kiezen	ankerpunten selecteren waarrond verschijnselen zijn gesitueerd en constante oriëntatie kiezen voor kaart	bij gesymboliseerde voorstelling een reëel beeld oproepen en zelf verschijnselen groeperen om ruimtelijke relaties duidelijk te maken	voorstelling met eigenschappen traditionele geografische kaart is de beste probleemrepresentatie voor ruimtelijke problemen

houdelijke kennis zo te structureren dat een degelijke oplossingsprocedure voor het probleem wordt gevonden (De Corte, 1981). Ook wij hebben aangenomen dat voor het goed begrijpen en oplossen van ruimtelijke problemen men een efficiënte ruimtelijke representatie moet kunnen opbouwen. Dit is slechts mogelijk wanneer men inzicht heeft in de begrippen die aan de basis liggen van zulke voorstelling. We hebben de hypothese gesteld dat een cognitieve voorstelling gelijkend op een geografische kaart bijzonder geschikt is om een ruimtelijk probleem voor te stellen. De geografische kaart wordt opgebouwd met vier concepten: perspectief, schaal, oriëntatie, symboolvoorstelling. Dit komt neer op het inzicht bijbrengen in de basistransformaties: veranderen van het eigen gezichtspunt naar een ander perspectief, het oriënteren van een omgeving ten opzichte van vaste punten, een onoverzichtelijke ruimte op schaal voorstellen, het selecteren en abstraheren van realistische waarnemingen naar een symbolische voorstelling. Wanneer men inzicht heeft in deze vier concepten kan men flexibel een ruimtelijke voorstelling opbouwen aangepast aan het ruimtelijk probleem dat zich stelt.

De vier ruimtelijke basisconcepten (perspectief, oriëntatie, schaal, symboolvoorstelling) zijn al vroeger in de literatuur vernoemd. Zo analyseert Gerber (1981) het begrip van de elementen van een kaart bij kinderen tussen 6 en 8 jaar. Hij vertrekt daarbij van de idee dat ontwikkelende leerlingen eerst de sleutelelementen van een kaart moeten verstaan voor zij de kaart als geheel kunnen begrijpen. De sleutelelementen van een kaart zijn voor hem: "a) arrangements of objects representing their actual locations on the earth's surface, mobility on a spatial reference system; b) an orthogonal (plan) view; c) an element of proportion which is based on scale; and d) a language which codifies information via a range of signs" (Gerber, 1981, p. 128). Hij vermeldt dus achtereenvolgens oriëntatie, perspectief, schaal en symboolvoorstelling. Op basis van zijn experimenten bij 80 kinderen tussen 6 en 8 jaar besluit hij dat zij de minste moeite hebben met het vogelperspectief; hun inzicht in verhoudingen en het begrip schaal is ronduit slecht en ook hun definitie van wat een kaart is loopt mank. In de Australische hoek vinden we tevens het onderzoek van Wilson (1981). Hij deed onderzoek bij 8- tot 12-jarigen. Hij analyseerde o.a. vrije schetsen op basis van de beheersing van de vier basisconcepten. Deze proefgroep heeft de minste moeilijkheden met het innemen van een ander perspectief; zich oriënteren en het voorstellen van de realiteit in symbolen liggen iets moeilijker. Achteraf gezien blijkt de manipulatie van verhoudingen en dus het inzicht

in het begrip schaal de moeilijkste opgave.

2.2 *Het trapsgewijs aanleren van de ruimtelijke basisbegrippen*

Om een betere ruimtelijke voorstellings- en structureringsvaardigheid bij de leerlingen te creëren, moeten ze inzicht verkrijgen in de vier ruimtelijke begrippen die aan de basis liggen van het kaartbegrip. Voor het aanleren van de begrippen inspireren we ons op de trapsgewijze methode van Gal'perin. Volgens Gal'perin moet men zich concentreren op "het transformeren van de materiële handeling die de leerling aanvankelijk uitvoert in een mentale handeling" (Van Parreren & Carpay, 1980, p. 33). Een mentale handeling houdt in dat er transformaties worden uitgevoerd op een object dat niet aanwezig is. De handeling gebeurt dus volledig abstract in het geheugen met behulp van een mentale voorstelling. Dit moet in fasen geleerd worden.

Gal'perin legt nadruk op de oriënteringsfase die het aanleren van een materiële handeling moet voorafgaan. In deze fase vertelt men de leerlingen wat ze met het materiaal precies moeten doen en op welke aspecten ze moeten letten. Wanneer de leerlingen voldoende voorbereid zijn om de materiële handeling correct uit te voeren, hebben ze een volledige oriënteringsbasis. Wij veronderstellen dat 12-jarigen beschikken over de algemene oriënteringsbasis voor de verschillende materiële handelingen in verband met het ruimtelijk voorstellingsvermogen, nl. het kunnen observeren van de zichtbare omgeving en het onderscheiden van de verschillende voorkomende vormen aan de hand van grootte, kleur, relatieve ligging en functie.

Op het materieel nivau moeten de objecten op een analoge manier in de realiteit worden gemanipuleerd als later op het mentaal niveau noodzakelijk zal zijn. De overgang naar de verbale handeling is de eerste stap naar abstractie; een verbale formulering moet in de plaats van de realiteit komen; deze verwoording onthouden leerlingen immers beter en exacter; in plaats van deze verbalisering menen wij dat het begrijpen en werken met een topografische kaart een meer geschikte eerste abstractie van de ruimtelijke werkelijkheid is. Ruimtelijke informatie op een verbale manier voorstellen houdt steeds een belangrijke reductie in van de ruimtelijke dimensie. Het tweede abstractieniveau noemen wij het 'cartografische niveau'. Tenslotte moet de leerling op het mentaal of abstracte niveau de basistransformaties kunnen uitvoeren op gegevens die niet aanwezig en waarneembaar zijn.

Onze keuze voor het trapsgewijs aanleren van de ruimtelijke basisbegrippen staat niet alleen. Zoals reeds vermeld is er in Australië een concentratie van onderzoek in deze richting. We citeren: "It is more realistic to consider young developing minds understanding key elements of a map before they can understand the map as a whole" (Gerber, 1981, p. 128). Uit dezelfde hoek: "Teachers should use concrete learning experiences based on the four elements of a map – arrangement, proportion, orthogonal view, abstraction, i.e. cartographic language – to foster the development of pupils' map reasoning ... The conclusions of this study imply that children need considerable and varied experiences, both inside and outside the classroom, with the four elements of a map to foster map reasoning development ... map work which cater for pupils at an abstract reasoning level should not be introduced until late in the years of elementary school, or even delayed until secondary school" (Wilson, 1981, p. 186). Ook hier vinden we hetzelfde idee: eerst de vier basisbegrippen grondig aanleren vooraleer de leerlingen met het globale kaartbegrip te confronteren. Dit moet gebeuren via leerervaringen zowel in de klas als er buiten. Het confronteren met een abstract niveau van ruimtelijk redeneren kan slechts in de laatste jaren van de basisschool en kan beter uitgesteld worden tot het begin van het secundair onderwijs. We onderstrepen het standpunt van Wilson dat het abstracte niveau van redeneren een extra nadruk moet krijgen. Inderdaad, de meeste cursussen en onderwijsvoorschriften in verband met kaartbegrip stoppen vooraleer dit abstractieniveau is bereikt. Dit probleem wordt recent ook beschreven in ons taalgebied: "Toch blijkt dat veel elf- en twaalfjarigen in de praktijk nog moeite hebben met het kaartgebruik, bijvoorbeeld als middel bij het vinden van een weg. Zij weten echter heel goed wat de verschillende kaartsymbolen betekenen en ook hoe het register gebruikt moet worden. Eén van de oorzaken hiervan zou wel eens kunnen zijn dat niet zozeer het technische kaartlezen tekortschiet als wel dat wat we met een algemene term de "ruimtelijke oriëntatie" noemen. Heel een-

voudig gezegd is ruimtelijke oriëntatie het vermogen om met of zonder hulpmiddelen iets te kunnen zeggen over de plaats die men op een bepaald moment in een grotere ruimte inneemt. We spreken van ruimtelijke oriëntatie wanneer iemand die tijdens een wandeling verdwaald is betrekkelijk gemakkelijk de weg terug kan vinden, bijvoorbeeld door een goed gebruik van aanwezige oriënteringspunten" (Aarts & Van den Bosch, 1984, p. 72). De leerlingen moeten dus verder gebracht worden dan het louter lezen van een kaart, ze moeten leren de symbolische vorm en constructie van een kaart aan de reële wereld te koppelen. Omgekeerd moeten ze ook van een reëel probleem een voorstelling kunnen maken die rekening houdt met de eisen van de vier ruimtelijke begrippen die een bruikbare voorstelling ondersteunen.

3 *Verloop van de experimentele lessen*

3.1 *De onderwijsvoorschriften*

De fundamentele doelstelling van deze lessen is dat we de leerlingen de vier ruimtelijke begrippen aan de basis van het kaartbegrip expliciet willen aanleren. Daarbij willen we verder gaan dan het niveau van het concrete en het technische kaartlezen, door aandacht te besteden aan het abstracte niveau. Tabel 1 geeft een samenvatting van de inhoud van de begrippen op elk abstractieniveau.

Perspectief

In de geografie betekent perspectief: de vaardigheid om de reële wereld zo op een tekening of kaart af te beelden, dat zij op het oog dezelfde indruk maakt als de reële wereld zelf, uit een gegeven standpunt gezien. Dit gegeven standpunt kan zowel de grond zijn, als een toren of een vliegtuig. De leerlingen moeten op materieel niveau begrijpen dat de wereld er anders uitziet naargelang het standpunt dat men inneemt. De omgeving van de school b.v. ziet er anders uit vanop de speelplaats, vanop de nabijgelegen kerktoren of vanuit een vliegtuig. Op cartografisch niveau maakt men zeer duidelijk een keuze voor het vogelperspectief: "perspectiefverandering: de gewone horizontale kijkrichting met een bewegende positie op de grond, wordt vervangen door een verticale, vaste kijkrichting met een positie vanuit de lucht" (Antrop, 1986, p. 130). Een topografische kaart is een bovenaanzicht van een gekozen gebied sterk gelijkend op een luchtfoto en dus op wat een piloot ziet. Op abstract niveau moeten de leerlingen een omgeving die op kaart is voorgesteld kunnen visualiseren in een beeld dat de omgeving voorstelt vanuit een normaal standpunt op de grond. We kunnen leerlingen b.v. vragen de juiste foto, genomen vanaf de grond, aan een bepaald gebied op de kaart te koppelen. Omgekeerd moeten ze ook een omgeving die ze nog nooit uit de lucht gezien hebben in een kaart met verticaal perspectief kunnen schetsen. Denken we b.v. aan het voorstellen op kaart van een omgeving waarin zich een verhaal uit een roman afspeelt. Hun eigen woon- en schoolomgeving kennen de meeste leerlingen alleen via het doorkruisen van het gebied op de grond. Ze moeten daarvan een kaart vanuit vogelperspectief kunnen schetsen.

Schaal

In geografische context gaat het om een verkleining van de reële wereld om deze te kunnen voorstellen. De geografische schaal geeft dus aan volgens welke verhouding de werkelijkheid op een kaart is verkleind. Op materieel niveau moeten de leerlingen begrijpen dat de verkleining van elke afstand op één geografische voorstelling volgens dezelfde verhouding is gebeurd. Dit in tegenstelling met wat we zien op terrein: wat dichtbij staat is groot en naar de horizon toe is er een schaalverkleining. De leerling moet b.v. beseffen dat een huis dat vlakbij staat op een geografische kaart dezelfde grootte zal krijgen als hetzelfde huis dat men in de verte ziet staan. Dit is een eerste stap in de abstractie van het onmiddellijk zichtbare. Op cartografisch niveau moet de leerling leren werken met de kaartschaal: de gehele kaart is een verkleining van een werkelijke omgeving volgens één en dezelfde schaal. Hierbij verkiezen we eerst het werken met de lijnschaal en pas daarna met de breukschaal. De lijnschaal staat dichter bij de realiteit: b.v. het lijntje van 10 cm is in werkelijkheid 10 km. De lijnschaal laat ook een gemakkelijke vergelijking toe met een gemeten afstand op de kaart: b.v. 2 cm tussen die twee dorpen is dus 2 km. Het werken met de breukschaal staat verder weg van de werkelijkheid: de omzetting van de gegevens 1/100 000 en 2 cm naar 2 km is veel moeilijker voor de leerlingen. Niet in het minst omdat leerlingen zich

eenvoudig misrekenen (al die nulletjes) of fouten maken bij de overgang van cm naar m en km. Nochtans vinden wij dat er moet opgepast worden met omkaderde stellingen als: "door het vermenigvuldigen van de gemeten afstand met de noemer van de breukschaal, vinden we de werkelijke afstand". Leerlingen die het concept schaal niet begrijpen zijn nogal gemakkelijk geneigd zulke regels uit het hoofd te leren. Wanneer ze die dan toepassen en zich misrekenen, zijn ze niet in staat in te zien dat hun bekomen resultaat onmogelijk is. Op abstract niveau moeten de leerlingen bij het bekijken van verschillende groot- en kleinschalige kaarten een idee hebben van de verkleining van de werkelijkheid en daaraan een reëel beeld kunnen koppelen. Omgekeerd moeten zij voor het maken van een ruimtelijke voorstelling een gepaste verkleining kunnen kiezen. Ze moeten b.v. kunnen voorzien bij aanvang dat bij het schetsen van een kaart van een omgeving het geheel op het blad papier kan getekend worden. Bovendien moeten ze de reële verhoudingen respecteren, ze moeten bewust zijn van de afstandsvervormingen die kunnen optreden. Dit laatste is wel een heel zware opgave.

Oriëntatie

Op materieel niveau betekent dit dat de leerlingen hun standplaats of bepaalde landschapskenmerken kunnen situeren in relatie met vaste herkenningspunten. Op dit niveau gaat het meestal om zichtbare punten. Het is belangrijk er de leerlingen op te wijzen dat het vaste punten moeten zijn. Mobiele herkenningspunten zijn niet geschikt als ankerpunt. Men kan zich b.v. niet efficiënt situeren in een stad t.o.v. de auto van de ijscoman; deze verplaatst zich immers regelmatig en dan kan men je via het herkenningspunt niet meer terugvinden. Op cartografisch niveau gebruikt men de kompasrichtingen om de kaart te oriënteren. De overeenkomst bestaat erin dat wanneer er geen aanduiding is de bovenkant samenvalt met het noorden. Het werken met een kompas laat de leerlingen beseffen dat het noorden inderdaad een vaste richting is om zich te oriënteren. De magnetische inclinatie mag maar in tweede instantie aandacht krijgen. Ook het gebruik van een roosterstructuur om zich te oriënteren op een kaart kan op dit niveau aan bod komen. Vertrekkende vanuit een eenvoudige rooster op het eigen stadsplan en het gebruik van een register van de plaatsnamen, kan men uiteindelijk ook de lengte- en breedteligging van een plaats leren bepalen en gebruiken. Op abstract niveau moeten de leerlingen bij het bekijken van een kaart een aantal ankerpunten leren selecteren, waarrond de andere verschijnselen gemakkelijk kunnen gesitueerd worden. Bij het zelf maken van een kaart moeten de leerlingen in functie van de voor te stellen aspecten en de gebruikers van de kaartschets een aantal vaste, herkenbare of gekende plaatsen aanduiden. De richtingen tussen deze punten moeten zo goed mogelijk met de werkelijke richtingen overeenstemmen. Ze moeten beseffen dat er systematische afwijkingen kunnen ontstaan. Net als bij het schaalbegrip is ook dit geen eenvoudige opdracht.

Symboolvoorstelling

Op materieel niveau moeten de leerlingen begrijpen dat niet alle elementen van het landschap volgens schaal kunnen verkleind worden en dan nog herkenbaar voorgesteld worden op de kaart. In plaats daarvan gebruikt men vereenvoudigde tekentjes. Gelijkaardige landschappelijke elementen worden door eenzelfde tekentje voorgesteld. Daarvoor moeten de leerlingen gelijkwaardige verschijnselen leren classificeren onder één noemer. Hier kunnen ze leren het onderscheid maken tussen menselijke en natuurlijke elementen in het landschap. Op cartografisch niveau leren de leerlingen de legende hanteren: dit is een lijst waarin men terugvindt wat een tekentje voorstelt. Het is een lijst van afspraken. Voor de topografische kaarten zijn de symbolen en kleuren gekozen dat we als het ware de wereld uit de lucht zien. Op dit niveau hoort ook het aanbrengen van het verschil tussen zichtbare en niet-zichtbare landschapselementen. Op abstract niveau moet men aan de gesymboliseerde voorstelling van de omgeving een reëel beeld kunnen koppelen. Tevens moet men bij het bekijken van een kaart een verdere abstractie kunnen doorvoeren: gelijkaardige verschijnselen ontdekken, classificeren en benoemen. Bij het zelf ontwerpen van een kaart moeten gelijkaardige verschijnselen via eenzelfde symbool voorgesteld worden. Men moet op die manier relaties duidelijk maken die niet onmiddellijk zichtbaar zijn. Hier bevinden we ons al in de grenszone tussen het louter voorstellen van een ruimtelijke omge-

ving en de geografische analyse van een gebied. Ons studiedomein stopt hier, terwijl het onderzoeksgebied van de Nederlander Van der Schee hier begint. Hij deed empirisch onderzoek naar het gebruik van geografische denkvaardigheden bij het analyseren van kaarten door leerlingen uit het vierde leerjaar van het V.W.O. (Van der Schee, 1987).

Kaartbegrip

Naast het verwerven van inzicht in al deze begrippen afzonderlijk, moet er ook nog aandacht besteed worden aan de integratie ervan tot een volwaardig kaartbegrip. Op materieel niveau betekent dit dat de leerlingen aandacht leren hebben voor de plaats waar zij zich bevinden en ook voor de plaats waar iets anders gelokaliseerd is of voor de omgeving waar een verschijnsel optreedt. Op cartografisch niveau leren ze dat een kaart een geschikt hulpmiddel is om het waar van de dingen te registreren en op te zoeken. De voorgestelde kaart is een verkleinde weergave van het landschap, vanuit de lucht gezien, georiënteerd met de windrichtingen, waarop landschapselementen door overeengekomen tekens worden voorgesteld. Op abstract niveau moeten ze ervan doordrongen zijn dat een voorstelling die gelijkt op een geografische kaart zeer bruikbaar is om problemen met een ruimtelijke dimensie voor te stellen en beseffen dat het vinden van een oplossing er aanzienlijk gemakkelijker door wordt.

3.2 *Concrete opbouw en invulling van het onderwijsexperiment .*

Een Vlaamse basisschool werd zo gastvrij gevonden om ons onderdak en medewerking te verlenen voor de experimentele lessen. Deze werden georganiseerd in twee parallelle klassen van het zesde leerjaar, met zowel jongens als meisjes. We beschikten over een geschikte doelgroep, namelijk leerlingen die aan het einde van het basisonderwijs gekomen zijn en dicht bij het begin van het secundair onderwijs. Wij en vele andere auteurs beschouwen deze leeftijd en periode als het meest geschikt voor de aanvang van een systematisch onderwijs in het kaartbegrip. De meeste leerlingen zijn dan klaar om de stap naar het abstracte niveau van redeneren en een coherent kaartbegrip te maken.

Volgens de directeur en onderwijzers waren de twee klassen globaal van hetzelfde niveau. Dit hebben we voor wat het kaartbegrip betreft wel nagegaan via een voortoets. In zowel de experimentele klas als de controleklas is dezelfde voortoets afgenomen om te controleren of er grote verschillen in voorkennis aanwezig zijn.

Vervolgens beschikten we in de experimentele klas over drie lesuren. In de controleklas werden er ondertussen geen aardrijkskundelessen gegeven. Onder deze omstandigheden moesten we wel een keuze maken wat de lesinhoud betreft. In drie lesuren kunnen niet de vier ruimtelijke concepten en het kaartbegrip op de verschillende niveaus ten gronde aangebracht en geoefend worden. We hebben getracht bij de leerlingen toch de vier begrippen expliciet aan te brengen vertrekkend vanuit het materiële naar het abstracte niveau. Dit laatste bleef beperkt tot het meegeven aan de leerlingen van inzicht in de basisbegrippen, om te weten aan welke aspecten ze aandacht moeten schenken bij het bekijken en zelf maken van een ruimtelijke voorstelling. Op het kunnen visualiseren van een kaartbeeld naar een reëel beeld zijn we praktisch niet ingegaan. We hebben de begrippen met het minimum aan uitbreidingen en oefening aangebracht. Sommige inzichten kunnen we zeer vlot bijbrengen omdat de leerlingen natuurlijk wel enige voorkennis hebben, maar van een coherent kaartbegrip is geen sprake. We kozen voor een tamelijk sturende onderwijsmethode. Tijdsgebrek liet ons niet toe het zelfontdekkend leren veel plaats te geven. Bovendien menen we dat bij het aanbrengen van basisbegrippen een sturende methode die expliciet het belangrijke aanwijst en laat oefenen meer rendabel is.

Nadien hebben beide klassen eenzelfde natoets ingevuld. Daarmee kunnen we controleren of de experimentele klas door onze drie lessen de flarden overgebleven voorkennis in verband met het kaartbegrip heeft kunnen reorganiseren in een samenhangend geheel en een abstracter redeneringsniveau heeft verworven.

Onze onderwijsvoorschriften en lessen hebben we vervolgens doorgegeven aan drie leerkrachten die in de gewone klaspraktijk staan. Ze hebben de lessen gegeven bij aanvang van het eerste jaar secundair onderwijs bij het leerplanonderdeel over landschap en kaart. In elke school is hetzelfde scenario gevolgd: de voortoets in de experimentele klas-

Tabel 2 *Resultaten van de experimentele klas uit het basisonderwijs (gemiddelde scores en score op Mann-Whitney test)*

	voortoets (max = 61)		natoets (max = 72)	
	experiment	controle	experiment	controle
jongens	44.7	42.8	42.3	34.0
meisjes	38.1	43.5	38.4	32.1
Totaal	41.0	43.3	40.1	32.8
Mann- μ Whitney α 32 deelnemers	103.5 0.3608		59.0 0.0085	

sen en de controleklassen, de lessen op basis van onze onderwijsvoorschriften aan de experimentele klassen en de traditionele lessen over landschap en kaart aan de controleklassen, de natoets in alle klassen.

4 *Resultaten van het onderwijsexperiment*

4.1 *Impact van de lessen op de experimentele klas van het zesde leerjaar basisonderwijs*

Verschillende opdrachten uit de voortoets en natoets zijn ontleend aan of geïnspireerd op proeven uit het constaterend onderzoek. We peilen telkens naar het inzicht in elk van de vier ruimtelijke basisbegrippen afzonderlijk, maar ook naar het geïntegreerd kaartbegrip.

Om de resultaten van de experimentele klas en de controleklas te kunnen vergelijken, berekenen we in eerste instantie de gemiddelde scores op de voortoets en de natoets, zowel voor de totale klas als voor de jongens- en meisjesgroep afzonderlijk (Tabel 2). Dit laatste is geïnspireerd door de vaststelling in vroeger onderzoek dat er een belangrijk verschil is in het inzicht in ruimtelijke begrippen. Voor de voortoets leiden we uit de gemiddelde totaalscores af dat de controleklas (43.3) iets beter presteert dan de experimentele klas (41.0). Deze laatste heeft een lagere score door de zwakke prestatie van de meisjes (38.1). Merk op dat in de controleklas de meisjes (43.5) net iets betere resultaten halen dan de jongens (42.8). Voor de natoets stellen we vast dat de gemiddelde totaalscore van de experimentele klas (40.1) nu een stuk boven die van de controleklas ligt (32.8). De zwakke groep meisjes uit de experimentele klas (38.4) blijven achter op de jongens (42.3), maar gaan relatief heel veel vooruit ten opzichte van de controleklas en ze maken een inhaalbeweging op hun mannelijke leeftijdsgenoten.

We hebben het verschil in de resultaten van de twee groepen statistisch getest. We kiezen voor de 'Mann-Whitney'-test die geschikt is voor onze ordinale gegevens over twee onafhankelijke groepen. Het is een nietparametrische test die weinig assumpties maakt over de aard van de data. De nulhypothese (experimentele klas en controleklas hebben dezelfde verdeling) kan in de voortoets niet worden verworpen, de testscore ($U = 103.5$) komt daarvoor overeen met een onaanvaardbaar significantieniveau ($\alpha = 0.3608$). In de natoets scoren de klassen wel significant verschillend, de waarschijnlijkheid dat we de nulhypothese onterecht verwerpen is zeer laag ($\alpha = 0.0085$ voor $U = 59$). Samenvattend kunnen we het volgende besluiten. Via de voortoets krijgen we zicht op de uitgangspositie van de twee klassen. Ze lossen de opgaven met ruimtelijke problemen op een quasi gelijke manier op. Uit de resultaten van de natoets blijkt dat de klas

die de experimentele lessen heeft gevolgd, beter presteert dan de controleklas. Op de eerste plaats betekent dit dat de vaardigheid om de ruimte voor te stellen en te structureren onderwijsbaar is. Op de tweede plaats dragen deze positieve resultaten bij tot de bevestiging van onze hypothese die aan de basis ligt van de onderwijsvoorschriften voor de experimentele lessen. Een ruimtelijke probleemrepresentatie die de eigenschappen heeft van een klassieke geografische kaart is dus een goede mentale ruimtelijke probleemrepresentatie die de kans op het vinden van een efficiënte oplossing voor een ruimtelijk probleem verhoogt.

4.2 *Het effect van de lessen in de gewone klaspraktijk*

In het onderwijsexperiment in het zesde leerjaar van een basisschool werden de lessen door de onderzoeker zelf gegeven. Dit in gekozen omstandigheden die ongetwijfeld afwijken van de reële school- en klassituatie waarin de meeste leerkrachten van het eerste jaar secundair onderwijs hun lessen in verband met het kaartbegrip geven. Vooraleer we onze resultaten kunnen generaliseren, moeten we ze in normale klassen en situaties toetsen. Dit kan uiteraard slechts door de lessen te geven aan meer klassen en verschillende groepen leerlingen. Het is niet zo eenvoudig om geschikte medewerkers op te sporen en bereid te vinden tot deelname aan een onderwijsexperiment. In België bestaat daar geen institutionele structuur voor en zijn we dus aangewezen op welwillendheid. Uiteindelijk kon het lesexperiment in het eerste jaar secundair onderwijs van drie scholen worden georganiseerd. Een belangrijk verschil met het eerste onderwijsexperiment is dat nu ook de controleklassen lessen krijgen over het kaartbegrip. De verschillen in de resultaten van de natoets tussen experimentele en controleklassen zullen hierdoor wellicht geringer zijn, dit omdat een aantal inzichten die wij belangrijk vinden ongetwijfeld ook in de traditionele lessen in zekere mate worden geoefend.

De analyse van de gemiddelde scores en testscores (Tabel 3) levert de volgende resultaten. We bekijken eerst de voortoets. In school 1 zijn de experimentele klas en de controleklas van vergelijkbaar niveau. In school 2 is de experimentele groep iets sterker dan de controlegroep. In school 3 zijn de groepen gelijkwaardig. Bestuderen we nu de natoets. De

Tabel 3 *Resultaten van het onderwijsexperiment in de gewone klaspraktijk (scores op Mann-Whitney test)*

	voortoets	natoets
school 1	46 deelnemende meisjes	
	μ = 240.5	μ = 135.0
	α = 0.5957	α = 0.0044
school 2	80 deelnemende jongens	
	μ = 547.5	μ = 534.0
	α = 0.0150	α = 0.0104
school 3	143 deelnemers (meisjes + jongens)	
	μ = 2120.0	μ = 1963.0
	α = 0.3395	α = 0.1060

experimentele klas in school 1 scoort duidelijk beter dan de controleklas. Ook in school 2 behalen de experimentele klassen een duidelijk beter resultaat dan de controleklassen, zij hadden natuurlijk wel al een voorsprong. Voor school 3 is de verbetering slechts zwak vast te stellen. We moeten echter rekening houden dat het een grote gevarieerde groep is van zwakkere en sterkere leerlingen, jongens en meisjes.

Deze resultaten willen we nog niet te absoluut formuleren. Allereerst gaat het nog steeds om een beperkt aantal klassen. Verder zijn deze op een eerder toevallige wijze in het experiment terecht gekomen. Een meer gecontroleerde samenstelling van de groepen is noodzakelijk. Tijd, middelen en medewerkers moeten daarvoor voor handen zijn. Tevens hebben we slechts in beperkte mate zicht gekregen op wat de leerkracht precies heeft aangeleerd in de lessen en op welke manier. De mate waarin de leerkrachten de nieuwe ideeën hebben verworven en de mate waarin die daadwerkelijk in de lessen zijn ingebouwd werden niet expliciet gemeten.

5 *Naar de huidige klaspraktijk en het onderwijsbeleid . . .*

Via toetsvragen die gesteld zijn in het eerste jaar secundair onderwijs in de loop van een volledig schooljaar in 149 scholen van de provincie Antwerpen (Goossens, 1988), hebben we de huidige klaspraktijk in verband met het

kaartbegrip geanalyseerd. De toetsvragen verraden als het ware wat de leerkracht belangrijk vindt. Voor de inhoudelijke doorlichting van de toetsvragen gebruiken we ons eigen interpretatiekader in verband met het kaartbegrip. We zoeken in elk van de 1577 toetsvragen over landschap en kaart welk ruimtelijk begrip moet worden toegepast en op welk abstractieniveau het probleem wordt gesteld. Daaruit kunnen we besluiten dat het begrip oriëntatie het meest wordt geëvalueerd, gevolgd door schaal en legende. Het globale kaartbegrip wordt relatief weinig getoetst en perspectief komt praktisch niet aan bod. Men ondervraagt overwegend op het cartografische niveau, het materieel en abstract redeneringsvermogen krijgen veel minder aandacht.

Met de leerplanonderrichtingen in verband met het kaartbegrip kunnen we ons verzoenen (Katholiek Nationaal Onderwijs, 1983). In het leerplan komen de basisbegrippen voldoende aan bod. Potentieel zijn er voldoende aanknopingspunten om onze aanpak te realiseren. Maar vanuit leerpsychologisch standpunt is een herschikking van de structuur en timing nodig, alsook aandacht voor de abstractiegraad van het redeneringsniveau. Wat betreft het abstractieniveau worden immers vooral het materiële en cartografische niveau benadrukt.

Om de opgedane kennis efficiënt te kunnen hanteren in het dagelijkse leven moet een hoger abstractieniveau nagestreefd worden: van identificeren naar beeldvorming. De leerlingen moeten de ruimtelijke demensie in zeer uiteenlopende problemen leren onderkennen en beoordelen, ze moeten aan de geleerde kaartvoorstelling een reëel beeld leren koppelen en gelijkaardige ruimtelijke verschijnselen onder één noemer classificeren. Deze uitdieping van het kaartbegrip kan blijven aan bod komen, ook in de latere jaren. Zowel gedurende de regionaal geografische benadering in de eerste vier jaren, als de thematisch geografische benadering in de laatste twee jaren van het secundair onderwijs.

Literatuur

Aarts, J. & H. van den Bosch, *Aardrijkskunde op de basisschool. Bouwstenen voor het onderwijsleerplan en de klaspraktijk*. Amsterdam: Van Goor Zonen, 1984.

Antrop M., Aardrijkskunde. Wat leren de kinderen op de basisschool. In: *Onderwijswegwijzer*, p. 111-142. Leuven: Wolters-Samson, 1986.

Corte, E. De, *Bijdragen tot de studie van onderwijsprocessen*. Leuven: Acco, 1981.

Gerber, R., Young Children's Understanding of the Elements of Maps. *Teaching Geography*, 1981, vol. 6, nr. 3, 128-133.

Goossens, M., Toetsvragen in het eerste jaar secundair onderwijs. *De Aardrijkskunde*, 1988, nr. 1, 29-55.

Katholiek Nationaal Onderwijs, *Leerplannen voor de eerste graad, aardrijkskunde*. Brussel: LICAP, 1983.

Parreren, C. van & J. Carpay, *Sovjetpsychologen over onderwijs en cognitieve ontwikkeling*. Groningen: Wolters-Noordhoff, 1980.

Schee, J. van der, Kijk op kaarten. *Nederlandse Geografische Studies*, 1987, nr. 36

Verhetsel, A., *De wereld in ons hoofd. Een onderwijsgeografisch onderzoek naar het ruimtelijk voorstellings- en structureringsvermogen*. Leuven: Katholieke Universiteit, Instituut voor Sociale en Economische Geografie, 1989.

Wilson, P., The map reasoning development of eight, ten and twelve year old pupils as revealed in free recall sketch maps. In: P. Wilson, R. Gerber & J. Fien (Eds.), *Research in Geographical Education*, pp. 143-199. Brisbane: AGERA, 1981.

Curriculum vitae

A. Verhetsel studeerde aardrijkskunde aan de K.U. Leuven (1978-1982). Nadien was ze als assistente verbonden aan het Instituut voor Sociale en Economische Geografie (1982-1987). In 1985 behaalde ze het Postgraduaat Getuigschrift in de Didactiek. In 1989 promoveerde ze aan de K.U. Leuven tot Doctor in de Wetenschappen, groep Aardrijkskunde met het proefschrift: "De wereld in ons hoofd. Een onderwijsgeografisch onderzoek naar het ruimtelijke voorstellings- en structureringsvermogen". Promotor was Prof.Dr. M. Goossens, copromotor Prof.Dr. E. De Corte. Momenteel is ze eerstaanwezende assistente aan de Faculteit Toegepaste Economische Wetenschappen van de UFSIA.

Adres: Universitaire Faculteiten St. Ignatius, Prinsstraat 13, B 2000 Antwerpen (België)

Manuscript aanvaard 18-1-'90

Summary

Verhetsel, A. 'The world in our head. An experimental tuiton program in spatial thought.' *Pedagogische Studiën*, 1990, *67*, 261-270

We are confronted with spatial problems in many everyday situations such as when we try to find the way to a new shop ar a friend's new house or when we try to find an alternative route to avoid traffic jams in the city. We are also living in an age of rapid and abundant diffusion of information from all corners of the world, altough we are only rarely conscious of this. Despite this plethora of information, many of us have a hazy knowledge of geography. Here you find the design, the organisation and the conclusions of an experimental tuiton program in 'spatial thought'. When twelve year old children get insight into the basic spatial concepts such as the ability to manipulate scale, perspective, orientation and symbols, they are able to construct mental spatial representations (cognitive maps) which allows them to find good solutions for spatial problems on a geographical size.

Discussie

Rekenmethoden-onderzoek op de proef gesteld

M. VAN DEN HEUVEL-PANHUIZEN
Vakgroep Onderzoek Wiskundeonderwijs en Onderwijscomputercentrum (OW & OC), Rijksuniversiteit Utrecht

Samenvatting

Een de laatste tijd veel gehoorde vraag is, of de gebruikte methode er wel toe doet. Harskamp c.s. beantwoorden naar aanleiding van hun implementatie- en effectonderzoek van reken-wiskundemethoden op de basisschool deze vraag ontkennend. Qua leerstof en qua leerresultaten blijken de methoden slechts marginaal te verschillen. In deze bijdrage wordt nagegaan in hoeverre deze conclusies en de bijbehorende aanbevelingen wel met recht worden verkondigd. Uit de analyse komt naar voren dat de basis ervoor maar zwak is.

1 *Inleiding*

Onderzoek naar de implementatie van nieuwe reken-wiskundemethoden en naar het effect van deze methoden op de leerresultaten van leerlingen is nodig, maar moeilijk. Nodig is het, omdat om het onderwijs en het rendement van het onderwijs te kunnen verbeteren, eerst bekend moet zijn hoe er in de praktijk met een bepaalde methode wordt gewerkt en wat dit oplevert aan leerresultaten. Moeilijk is het, omdat bij dit soort onderzoek zoveel variabelen in het spel zijn die bijna niet of zelfs helemaal niet te controleren zijn. Dit maakt het onderzoek niet onmogelijk, maar noopt wel tot de nodige voorzichtigheid met betrekking tot de conclusies en de aanbevelingen. Er dient dan ook kritisch te worden nagegaan of de conclusies en aanbevelingen wel met recht worden verkondigd. In deze bijdrage worden de conclusies die Harskamp en Suhre trekken naar aanleiding van hun implementatie- en effectonderzoek van reken-wiskundemethoden in het basisonderwijs aan zo'n kritische analyse onderworpen.

Via allerlei publikaties gericht op verschillende doelgroepen hebben Harskamp en Suhre laten weten:

- dat de methoden qua leerstof slechts heel weinig van elkaar verschillen (Harskamp & Suhre, 1987, 21 en 23; Suhre, 1988, 5),
- dat er een indicatie is, dat de (moderne) methoden worden gebruikt zoals bedoeld (Harskamp, 1988, 85)[1],
- dat er verschillen bestaan in de condities waaronder de methoden worden gebruikt (Harskamp & Suhre, 1988, 213-214), m.n. dat de moderne methoden veel op scholen worden gebruikt met leerlingen van een meer dan gemiddeld intelligentieniveau (Suhre, 1988, 6 en 5), dat de moderne methoden een uitdaging vormen voor intelligente kinderen en dat in heterogene klassen de stof voor zwakke leerlingen problemen op kan leveren (Harskamp & Suhre, 1987, 23),
- dat de verschillen in toetsprestaties en in rekenattitude tussen de leerlingen marginaal zijn (Harskamp & Suhre, 1987, 22 en 23; Harskamp & Suhre, 1988, 216; Harskamp, 1988, 86, 87 en 89),
- dat niet de methode, maar een aantal implementatiekenmerken – waaronder het percentage leerlingen waarmee de basisstof is behandeld – van invloed is op de leerprestaties (Harskamp & Suhre, 1987, 23; Harskamp & Suhre, 1988, 216; Harskamp, 1988, 88).

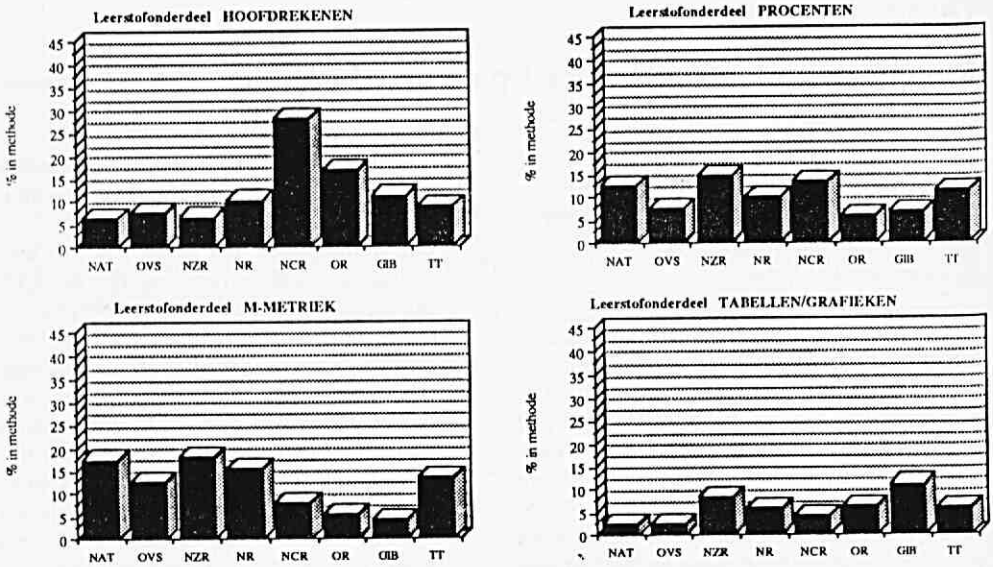
Zijn al deze conclusies wel houdbaar?

2 *De inhoud van de reken-wiskundemethoden*

De inhoudsanalyse van de methoden, die in het kader van het onderzoek van Harskamp en Suhre is uitgevoerd, omvat een leerstofanalyse en een didactische analyse.

2.1 *Leerstofanalyse*

Als eerste wordt ingegaan op de leerstofanalyse van de methoden en de conclusie dat de methoden wat betreft de relatieve aandacht voor de verschillende leerstofonderdelen weinig



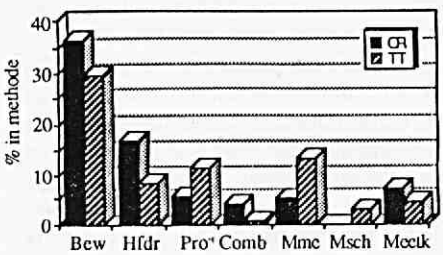
NAT=Naar Aanleg en Tempo; OVS=Op Veilig Spoor; NZR=Naar Zelfstandig Rekenen; NR=Nieuw Rekenen; NCR=Niveaucursus Rekenen; OR=Operatoir Rekenen; GIB=Getal in Beeld; TT=Taltaal

Figuur 1 *Percentage aandacht dat de verschillende methoden besteden aan bepaalde leerstofonderdelen (percentages overgenomen uit Harskamp, 1988, 39)*

van elkaar verschillen. De toegepaste analyseprocedure – per taak zijn drie punten te verdelen over een aantal vooraf onderscheiden leerstofcategorieën (Harskamp, 1988, 36-37) – is nogal zwak om als basis te dienen voor deze conclusie.

Een bezwaar is, dat er sprake is van een tamelijk grove indeling waardoor het niet altijd gemakkelijk is om uit te maken bij welke categorie een bepaalde opgave hoort. Maar ook zonder dit probleem is het nog de vraag of op grond van de gevonden percentages (o.c., 39) wel geconcludeerd kan worden dat "er (...) over het geheel genomen geen grote verschillen (zijn, mvdh) in het percentage leerstof dat per methode wordt besteed aan de verschillende leerstofonderdelen" (Harskamp, 1989, 3). Aan de hand van de door Harskamp gevonden percentages kan ook een andere conclusie worden getrokken. Namelijk, dat er met betrekking tot de aandacht die methoden besteden aan bepaalde leerstofonderdelen soms grote verschillen bestaan. Zie bijvoorbeeld de verschillen in relatieve aandacht die er tussen de methoden bestaan met betrekking tot de leerstofonderdelen hoofdrekenen, procenten, metriek en tabellen/grafieken (zie Figuur 1). Dit zijn allemaal leerstofonderdelen waarop volgens Harskamp vrijwel alle methoden veel nadruk leggen (o.c., 3).

Bij de conclusie, dat er tussen de methoden over het geheel genomen geen grote verschil-



Bew = Bewerkingen
Hfdr = Hoofdrekenen
Pro = Procenten
Comb = Combinatoriek
Mme = Meten – Metriek
Msch = Meten – Schatten
Meetk = Meetkunde

Figuur 2 *Percentage aandacht dat de methoden OR en TT besteden aan de bepaalde leerstofonderdelen (percentages overgenomen uit Harskamp, 1988, 39).*

len zijn in aandacht voor de leerstofonderdelen, moet bovendien worden betrokken, dat in de onderwijspraktijk bijna nooit sprake is van een keuze uit alle methoden. Er is meestal sprake van een vergelijking van twee of drie methoden. En dan kunnen de verschillen in relatieve aandacht voor de verschillende leerstofonderdelen ook groot zijn. Zoals bijvoorbeeld het geval is bij de methoden Operatoir Rekenen en Taltaal (zie Figuur 2).

Iets dat de vergelijking ook moeilijk maakt, is de breedte van het leerstofonderdeel bewerkingen. Dit leerstofonderdeel bevat zowel opgaven die betrekking hebben op de hoofdbewerkingen, als opgaven met breuken en opgaven met kommagetallen. Onder deze overkoepelende categorie kunnen veel verschillen schuilgaan. Het belangrijkste bezwaar tegen de uitgevoerde leerstofanalyse is dan ook, dat er geen inhoudelijke analyse van de leerstof heeft plaatsgevonden. Pas dan kan worden gezegd, of er wat de inhoud betreft tussen methoden verschillen bestaan.

Hoe verschillend de leerstofinhoud kan zijn, zelfs als het gaat om een zeer duidelijk omschreven leerstofonderdeel als het leren optellen en aftrekken, moge blijken uit het volgende voorbeeld dat betrekking heeft op de methoden Naar Zelfstandig Rekenen (NZR) en De Wereld in Getallen (WIG) (zie Van den Heuvel-Panhuizen, 1989). Bij NZR leren de kinderen in groep 3 optellen en aftrekken tot 20 en hieronder vallen ook de sommen waarbij het tiental wordt overschreden. Bij WIG daarentegen leren de kinderen in groep 3 al sommen maken tot 100, maar er worden nog geen sommen gemaakt waarbij sprake is van een tientaloverschrijding. Deze leerstofinhoudelijke verschillen zouden bij een louter kwantitatieve analyse, op de manier zoals Harskamp (1988, 36-37) die heeft toegepast, niet naar voren zijn gekomen. Alles zou dan onder de noemer van bewerkingen zijn gevallen[2].

2.2 *Didactische analyse*

Naast de leerstofanalyse heeft er in het onderzoek van Harskamp c.s. ook een didactische analyse van de methoden plaatsgevonden. Ondanks dat de methoden qua relatieve aandacht voor de onderscheiden leerstofonderdelen in grote mate met elkaar overeenstemmen, verschillen de methoden volgens Harskamp wèl in de vakinhoudelijke vormgeving van de leerstof (o.c., 1988, 83 en 15). Harskamp beschrijft uitgebreid de bevindingen van De Jong (1986) aangaande de aanwezigheid en de getrouwheid van Wiskobaskenmerken in methoden en gebruikt deze methodebeoordelingen om de bij zijn onderzoek betrokken methoden in te delen in modern en traditioneel.

Het gegeven dat De Jong bij de vakinhoudelijke beoordeling van de leerstof in tegenstelling tot Harskamp wel tot grote verschillen komt tussen de methoden, wordt door Harskamp geenszins als problematisch beschouwd. De Jongs typering van methoden is immers uitsluitend op de aanwezigheid en getrouwheid van Wiskobaskenmerken gebaseerd (Harskamp, 1988, 83). Bovendien hoeven volgens Harskamp vakinhoudelijke verschillen niet in te houden, dat de moderne methoden op andere leerstofonderdelen nadruk leggen dan de traditionele methoden (o.c., 9). Uit de door Harskamp (1988, 7-8) vermelde lijst van Wiskobaskenmerken valt echter op te maken dat deze opvatting niet houdbaar is. De Wiskobaskenmerken hebben niet alleen betrekking op de vormgeving van de leerstof, maar ook op de keuze voor een bepaalde leerstofinhoud. Vandaar ook dat het werk van Wiskobas is terug te vinden in de eindtermen van de basisschool (vgl. Treffers, 1989).

De didactische analyse die Harskamp c.s. zelf hebben uitgevoerd, betreft de vakdidactische kenmerken leerstofdifferentiatie en variatie in leerstof (Harskamp, 1988, 37).

Voor de *variatie in leerstof* is gebruik gemaakt van de eerder uitgevoerde leerstofanalyse. Aan de hand van de geïnventariseerde percentages per leerstofonderdeel en een aanvullende inventarisatie voor het leerstofonderdeel redactieopgaven, is de mate van variatie in de leerstof vastgesteld. Deze mate van variatie wordt berekend door "de percentages van redactieopgaven en hoofdrekenen en eigenschapsrekenen bij elkaar op te tellen. Dit percentage geeft aan welk deel van de totale leerstof besteed wordt aan andere activiteiten dan het inoefenen van algemene rekenvaardigheden" en "het percentage algemene rekenvaardigheden in 'kale' vorm is nu een restcategorie, te bepalen door de percentages voor redactieopgaven en hoofdrekenen en eigenschapsrekenen af te trekken van 100%" (o.c., 41)[3]. Dit is een operationalisatie die

veeleer betrekking heeft op de mate waarin een methode aandacht besteedt aan bepaalde leerstof (redactieopgaven, hoofdrekenen en eigenschapsrekenen) dan op de variatie in leerstof. Doordat er vanuit wordt gegaan dat het percentage redactieopgaven, hoofdrekenen, eigenschapsrekenen en de algemene rekenvaardigheden in kale vorm samen 100% vormen, is het bovendien een operationalisatie die uitgaat van een tamelijk beperkte kijk op de inhoud van het vak rekenen-wiskunde. Leerstofonderdelen als meten, tabellen/grafieken en meetkunde tellen hierbij blijkbaar niet mee. Ter zijde zij nog opgemerkt, dat bij de meting van de door de leerkrachten uitgevoerde leerstofvariatie een enigszins andere operationalisatie van de leerstofvariatie is gehanteerd. Er is gevraagd naar het percentage van de lestijd dat besteed wordt aan hoofdrekenen, redactieopgaven (of toepassingsproblemen uit reële situaties) en algemene rekenvaardigheden (o.c., 127)[4]. Het eigenschapsrekenen is hierbij als aparte categorie komen te vervallen en de algemene rekenvaardigheden (oefeningen op het gebied van cijferen, metriek stelsel, meten, breuken, procenten, enz.) bestaan volgens deze omschrijving blijkbaar niet meer noodzakelijkerwijs alleen uit kale opgaven.

Terug naar de leerstofvariatie in de methoden. De door Harskamp gekozen operationalisatie levert een beeld op van de mate van leerstofvariatie in de verschillende methoden, dat met name door de plaats die de methode Niveaucursus Rekenen (NCR) in de rij inneemt, niet zo erg herkenbaar is. De hoge score van NCR is bijna volledig in tegenspraak met bijvoorbeeld de beoordeling die De Jong (1986), 183) geeft van NCR: "in alle opzichten een mechanistische methode, die de niveaucursus-organistie centraal stelt ten koste van een actief, gedifferentieerd (...) en wiskundig rijk onderwijs". Harskamp zelf noemt de score van NCR ook opvallend (o.c., 41), maar ziet hierin klaarblijkelijk geen aanleiding om de gekozen operationalisatie te heroverwegen.

Een alternatief is bijvoorbeeld: per methode de leerstofonderdelen in volgorde van grootte zetten en dan vervolgens kijken naar het aantal verschillende leerstofonderdelen waaruit de eerste 60% of 80% (te beginnen bij het leerstofonderdeel met het hoogste percentage) van het totale leerstofaanbod bestaat. Als dit wordt toegepast op de door Harskamp bij de leerstofanalyse gevonden percentages (met alle gebreken vandien), zou dit leiden tot een ander beeld van de mate van leerstofvariatie in de methoden (zie Figuur 3). Een beeld dat waarschijnlijk meer overeenkomt met de werkelijke variatie in leerstof(onderdelen) in de methoden. NCR bijvoorbeeld blijkt hierin juist laag te scoren.

methode	aantal onderdelen in 60% van de leerstof	aantal onderdelen in 80% van de leerstof	'variatie in leerstof' Harskamp (1988)
NCR	Bew Hfd	Bew Hfd Pro	45%
NZR	Bew Mm	Bew Mm Pro Tab	30%
OVS	Bew Mm	Bew Mm Ml Pro - Hfd	32%
NAT	Bew Mm	Bew Mm Pro Ver Get	32%
NR	Bew Mm	Bew Mm Hfd Pro Get	38%
OR	Bew Hfd	Bew Hfd Mtk Tab Pro - Get	49%
GIB	Bew Get Hfd	Bew Get Hfd Tab Pro Eig	45%
TT	Bew Mm Pro	Bew Mm Pro Hfd Get Tab	46%

Bew=Bewerkingen; Hfd=Hoofdrekenen; Pro=Procenten; Mm=Meten-Metriek; Ml=Meten-Aflezen; Tab=Tabellen/Grafieken; Get=Getallen; Eig=Eigenschapsrekenen

Figuur 3 *Variatie in leerstof(onderdelen) in methoden.*

3 *De implementatie van reken-wiskundemethoden*

Via een schriftelijke bevraging van de leerkrachten is nagegaan in welke mate zij een aantal implementatiekenmerken toepassen. Geconcludeerd wordt, dat het feit dat de methoden, zoals verwacht, verschillen op de twee vakdidactische implementatiekenmerken een indicatie geeft dat de (moderne) methoden worden gebruikt zoals bedoeld (Harskamp, 1988, 85). Gemiddeld worden de moderne methoden zelfs iets beter geïmplementeerd dan de traditionele (Suhre, 1987, 353). Betekent dit nu, dat uit het onderzoek van Harskamp en Suhre blijkt dat de vernieuwende kenmerken – want daar gaat het bij implementatie toch om (vgl. Fullan & Pomfret, 1977) – van de realistische methoden inderdaad gerealiseerd worden in de praktijk? Nee. Immers, geen enkele van de vernieuwende kenmerken van de realistische methoden, zoals het gebruik van contexten, de aandacht voor modellen, de grote mate van eigen inbreng van de leerlingen, het interactieve karakter en de onderlinge samenhang van de verschillende leerstofonderdelen (zie Treffers & Goffree, 1985; Treffers & Feijs, 1987) is door Harskamp en Suhre onderzocht. In plaats daarvan is gekozen voor implementatiekenmerken "die door de leerkracht op betrouwbare wijze kunnen worden beantwoord" en "die vermoedelijk voorspellende waarde hebben ten opzichte van leerprestaties" (Harskamp, 1988, 27). Het interactieve karakter van het onderwijs, dat als eerste bij de *vakdidactische implementatiekenmerken* wordt genoemd (o.c., 15-16) en waarin alle kenmerken van de realistische didactiek worden ondergebracht (Harskamp, 1989, 4), is bij nader inzien dan ook niet onderzocht. Het zou te moeilijk zijn te operationaliseren door de te globale aanwijzingen in de methoden, het zou alleen op de introductie van nieuwe leerstof slaan en het zou geen voorspellende waarde hebben ten aanzien van de leerresultaten (Harskamp, 1988, 27).

Wat overblijft zijn twee andere vakdidactische implementatiekenmerken: variatie in de leerstof en leerstofdifferentiatie. Voor beide kenmerken is een significant verschil gevonden tussen de acht methoden (o.c., 56)[5]. Bovendien is er een samenhang gevonden tussen datgene wat de didactische analyse van de methoden hierover heeft opgeleverd en de bevraging van de leerkrachten (o.c., 56). Het zijn deze beide gegevens die een indicatie geven dat de (moderne) methoden worden gebruikt zoals bedoeld (o.c., 85). Gezien de gekozen operationalisaties is de basis voor deze conclusie zwak. Voor de bezwaren tegen de operationalisatie van de variatie in leerstof kan verwezen worden naar de opmerkingen die zijn gemaakt bij de didactische analyse van de methoden. Ten aanzien van de operationalisatie van de leerstofdifferentiatie moet worden opgemerkt, dat de bij de theoretische achtergronden wel genoemde (o.c., 17), voor de realistische didactiek kenmerkende, proces differentiatie later niet terug komt bij de bevraging van de leerkrachten. Omdat de aanwijzingen omtrent procesdifferentiatie in de methoden te summier zijn, worden geen grote verschillen in instructiegedrag verwacht en wordt alleen gevraagd naar de mate waarin verschillende leerlingen in een klas verschillende leerstof krijgen tijdens de lessen (o.c., 28). Een bijkomend probleem is dat de categorieën die gebruikt zijn bij de methodenanalyse niet dezelfde zijn als bij de bevraging van de leerkrachten (vgl. o.c., 37 en 127).

Naast de twee vakdidactische implementatiekenmerken omvat het onderzoek ook nog drie *algemene implementatiekenmerken*: het percentage leerlingen waarmee de basisstof is behandeld, het aantal lessen per week en het aantal minuten dat men per les bezig is met rekenen. Ten aanzien van de laatste twee kenmerken is geen significant verschil tussen de acht methoden gevonden (o.c., 55-56). Bij het percentage leerlingen dat de basisstof heeft gedaan, is wel een significant verschil gevonden tussen de acht methoden, maar niet tussen de traditionele en de moderne methoden (o.c., 54). Ook deze kenmerken zeggen in feite niets over de mate waarin een bepaalde vernieuwing wordt gerealiseerd. Een leerkracht kan op deze manier gemakkelijk een hoge implementatiescore halen – door bijvoorbeeld veel lessen per week te geven of door lange lessen te geven – zonder dat iets van de vernieuwing in praktijk wordt gebracht. De kenmerken zijn veeleer algemene onderwijscondities dan implementatiekenmerken. Er moet een bepaalde tijd aan onderwijs worden besteed, wil het effect sorteren.

Met betrekking tot het percentage leerlin-

gen waarmee de basisstof is behandeld, ligt het wat moeilijker. Het is de vraag of dit ook als een algemene onderwijsconditie kan worden beschouwd of dat het eerder een gevolg is van de gebruikte methode. In § 5.2 wordt hier nog op teruggekomen.

4 *De condities waaronder de methoden worden gebruikt*

Hoe zit het met de condities waaronder de methoden worden gebruikt? Voor de beantwoording van deze vraag zijn gegevens verzameld over drie achtergrondkenmerken van de leerkrachten en twee achtergrondkenmerken van de leerlingen.

Met betrekking tot de *leerkrachten* is nagegaan hoeveel jaren ervaring zij hebben met de methode, of er sprake is van een leerling- of van een leerstofgerichte attitude en of ze wel of niet een verzwaard wiskunde-examen in hun vooropleiding hebben gehad. Alleen bij het aantal jaren ervaring is een significant verschil gevonden tussen de acht methoden (o.c., 58). Harskamp voegt hier aan toe, dat leerkrachten van moderne methoden gemiddeld genomen minder jaren ervaring hebben met hun methode dan leerkrachten van traditionele methoden (o.c., 58). Uit niets blijkt echter dat dit ook is getoetst. Het lijkt erop, dat aan het gegeven dat er een significant verschil bestaat tussen de acht methoden tevens de conclusie is verbonden dat er een verschil bestaat tussen de moderne en de traditionele methoden. Dit is een onjuiste handelwijze die meer dan eens voorkomt[6].

Wat betreft de *leerlingen* heeft er aan het eind van de basisschool een meting van de intelligentie plaatsgevonden en is er een milieuscore vastgesteld. Voor de intelligentie is gebruik gemaakt van de ISI-test en de milieuscore is gebaseerd op het beroep van de vader (o.c., 48). Analyse van deze gegevens leidt tot de conclusie dat "voor de verwachting dat moderne methoden vaker in klassen worden gebruikt met een wat hoger intelligentieniveau (. . .) ondersteuning (is, mvdh) gevonden" en dat er "significante verschillen tussen methoden (zijn, mvdh) ten aanzien van de gemiddelde intelligentiescores van klassen" (o.c., 59). Er worden echter geen toetsingsgegevens vermeld waaruit kan worden afgeleid dat er ook een toetsing heeft plaatsgevonden die een significant verschil heeft opgeleverd tussen de twee soorten methoden (zie o.c., 57). Pas als dit wel is gebeurd, kan er een conclusie worden getrokken over de relatie tussen intelligentie en type methode. En dan niet de conclusie die Harskamp heeft geformuleerd, maar de conclusie dat aan het eind van de basisschool de gemiddelde intelligentiescore hoger is in klassen met een moderne methode dan in klassen met een traditionele methode (zie ook de reactie van Freudenthal, 1989)[7].

5 *Verschillen in leerprestaties*

De kern van het onderzoek wordt gevormd door de vraag of er verschillen zijn in leerprestaties tussen de methoden als rekening wordt gehouden met de omstandigheden waaronder de methoden worden gebruikt. Om de effecten van de methoden te meten is gebruik gemaakt van de CITO-eindtoets rekenen en de CITO-eindtoets tabellen & grafieken. Verder zijn aanvullende RION-toetsen ontwikkeld voor de leerstofonderdelen schatten van standaardmaten, meten, verhoudingen en combinatoriek. Ook is nog een vragenlijst gebruikt waarmee wordt gemeten welk belang kinderen hechten aan rekenen en of ze er plezier in hebben (Harskamp, 1988, 44). Er hebben twee analyses plaatsgevonden, een op leerling- en een op klasseniveau.

5.1 *Analyse op leerlingniveau*

Begonnen is met de analyse op leerlingniveau. Via covariantie-analyses is nagegaan of er – als er wordt gecorrigeerd voor intelligentie – verschillen zijn tussen de rekenmethoden ten aanzien van de rekenprestaties van de leerlingen.

Wat betreft het corrigeren voor intelligentie kunnen zich volgens Harskamp twee belemmeringen voordoen (o.c., 59). De eerste belemmering heeft te maken met de beïnvloedbaarheid van intelligentie door onderwijs. Omdat het hier gaat om de achteraf gemeten intelligentie, is het corrigeren voor intelligentie alleen toelaatbaar als het onderwijs dat de leerlingen hebben gehad geen invloed heeft op de prestaties op de intelligentietest op het eind van de basisschool. Op grond van de geraadpleegde literatuur (Detterman & Sternberg, 1982; Slavenburg, 1986) is Harskamp (1988, 20 en 59) inderdaad die mening

toegedaan. Daarmee is de eerste belemmering van de baan. Maar hoe gegrond is die mening?

Op de eerste plaats is het belangrijk, dat een onderscheid wordt gemaakt tussen de mogelijkheid om via speciale trainingen een duurzame verbetering van het intelligentieniveau te bewerkstelligen en de mogelijkheid dat een aantal jaren onderwijs leidt tot een verhoging van het intelligentieniveau (zie ook Freudenthal, 1989). Het is echter niet duidelijk waar de aangehaalde literatuur betrekking op heeft. Wat betreft de invloed van het onderwijs op het intelligentieniveau, kan tegenover de mening van Harskamp in ieder geval de mening van Lohman worden geplaatst. In een recente publikatie stelt Lohman (1989), dat nog te vaak wordt gedacht dat intelligentietesten de aangeboren bekwaamheid van de leerling meten. Volgens hem meten deze testen juist een van de belangrijkste uitkomsten van het onderwijs, namelijk de vaardigheid om datgene wat op school is geleerd over te dragen naar nieuwe situaties. Naast allerlei vaardigheden die dicht liggen bij wat de kinderen onderwezen is, de zogenoemde 'crystallized abilities', is ook de 'fluid ability' een produkt van onderwijs en ervaring. Een onderzoek dat betrekking had op verschillende taal/leesprogramma's toonde zelfs aan dat verschillende programma's verschillende scores bij deze twee soorten 'abilities' kunnen opleveren. Bij de erg gestructureerde programma's met veel 'drill and practice' waren de vorderingen bij de 'fluid ability' het laagst en die bij de vakspecifieke vaardigheden het hoogst, terwijl bij de minder gestructureerde en meer op probleem oplossen gerichte programma's het omgekeerde patroon was te zien. Of met de woorden van Lohman: "You get out a program what you put into it" (Lohman, 1989, 9).

De tweede belemmering heeft betrekking op een mogelijk verschillende uitwerking van methoden op de rekenprestaties van kinderen met een verschillende intelligentie. Daarom wordt door middel van multivariate variantieanalyse (met methode en intelligentie als onafhankelijke variabelen en de verschillende toetsscores als afhankelijke variabelen) eerst gekeken of er een interactie bestaat tussen methode en intelligentie op de toetsprestaties. Dit blijkt niet het geval te zijn (Harskamp, 1988, 60).

Na de eliminatie van de twee mogelijke belemmeringen is overgegaan tot de covariantieanalyse met intelligentie als covariaat. Eerst is er multivariaat getoetst met alle afhankelijke variabelen. Deze multivariate toetsing leverde een methode-effect op (o.c., 60). Onduidelijk is of het hier behalve de factor 'methode', ook de factor 'type methode' betreft. Toetsingsgegevens ontbreken.

De vervolgens uitgevoerde univariate covariantie-analyses leverden geen significant verschil op tussen de methoden ten aanzien van de totaalscores bij de CITO-toets rekenen en de CITO-toets tabellen & grafieken (o.c., 62) en leverden ook geen significant verschil op bij de RION-toets als geheel (o.c., 60). Alleen voor het onderdeel combinatoriek is een significant verschil gevonden tussen de acht methoden én tussen de twee typen methoden (o.c., 60).

Behalve het significante verschil bij het toetsonderdeel combinatoriek, meldt Harskamp ook dat er een opvallend hoge samenhang is ($r = .9$) tussen de relatieve aandacht van de acht methoden voor het leerstofonderdeel combinatoriek en de gemiddelde leerlingscore per methode op dat onderdeel (o.c., 60). Ook voor twee andere onderdelen van de RION-toets, de onderdelen verhoudingen en meten zijn tamelijk sterke positieve correlaties (resp. $r = .9$ en $r = .7$) gevonden tussen de relatieve aandacht voor het betreffende leerstofgebied in de acht methoden en de gemiddelde toetsresultaten van de leerlingen per methode (o.c., 86)[8]. Harskamp hecht aan deze uitkomst kennelijk niet zoveel belang. Zijn conclusie is: "Sommige methoden leiden tot iets betere leerlingresultaten op onderdelen van de toetsen dan andere methoden, maar *over het geheel genomen zijn er dus geen verschillen tussen de methoden* (curs., mvdh) op zowel algemene rekenvaardigheden (CITO-toets) als op toepassingen (RION-toets)" (o.c., 86).

In de gevonden samenhang tussen de mate van aandacht voor een leerstofonderdeel en de prestaties op dat onderdeel ziet Harskamp een bevestiging van de 'conclusie' van Walker en Schaffarzick (1974), dat methoden vooral zullen verschillen in leerresultaten bij leerstofonderdelen die verschillende nadruk in de methoden krijgen (Harskamp, 1988, 86). De redenering van Harskamp c.s. is ongeveer als volgt: er zijn weliswaar lichte verschillen tussen methoden in nadruk op bepaalde leerstof-

onderdelen, maar over het geheel genomen maakt het voor de rekenprestaties niet uit met wèlke methode de kinderen les krijgen (Harskamp en Suhre, 1987, 23).

Dat is nu precies waar Walker en Schaffarzick voor waarschuwen! Zij begonnen – zoals zij zelf zeggen – hun studie naar onderzoeken waarbij curricula worden vergeleken zoals er altijd wordt begonnen, naïef, met het zoeken naar gegevens die zouden kunnen wijzen op de superioriteit van moderne curricula ten opzichte van traditionele curricula. Wat ze vonden was geen superioriteit, maar analogie. De curricula deden het beter met betrekking tot die leerstofonderdelen in de toetsen die het meest overeenkomen met wat het curriculum biedt. Dit blijkt echer zo'n vanzelfsprekende uitkomst te zijn, dat er door niemand aandacht aan wordt besteed. Walker en Schaffarzick willen daarentegen juist, dat er meer aandacht wordt besteed aan de voor de hand liggende uitkomst, dat verschillende curricula leiden tot verschillen in prestaties (vgl. Lohman, 1989). Aan het eind vatten ze het nog eens kernachtig samen: curricula moeten gezien worden als middel – en naar het schijnt als een machtig middel – om de leerprestaties van leerlingen te sturen. Kennelijk hebben Harskamp c.s. de essentie van wat Walker en Schaffarzick te zeggen hebben over het hoofd gezien. Ook zij gaan nogal achteloos voorbij aan de vanzelfsprekende uitkomst dat de mate van aandacht voor bepaalde leerstofonderdelen samenhang vertoont met de prestaties op die leerstofonderdelen.

Behalve naar de leerprestaties is door Harskamp c.s. ook gekeken naar de invloed van de methode op het belang dat kinderen hechten aan rekenen en het plezier dat ze er in hebben. Ook nu weer was er geen sprake van een significant verschil (Harskamp, 1988, 62). Voorstanders van de "moderne (zogenoemde 'realistische') rekendidactiek" verwachten volgens Harskamp, dat moderne methoden behalve tot betere leerprestaties, ook tot een positievere houding van leerlingen ten opzichte van rekenen leiden (o.c., 11). Dit is een misverstand. Bij de aangehaalde auteurs (Treffers, 1982; Van den Heuvel-Panhuizen & Goffree, 1986) kan dit ook niet worden gevonden. Als voorstanders van de realistische didactiek het hebben over een betere attitude van leerlingen, dan hebben ze het op de eerste plaats over een *wiskundige attitude*, een onderzoeksgerichte instelling en wiskundige aanpak van problemen (vgl. Freudenthal, 1978; Streefland, 1980). Plezier eraan beleven kan er bij komen, maar daarin hoeven de kinderen niet te verschillen van de kinderen die mechanistisch rekenonderwijs krijgen. De resultaten van een hypothesen-onderzoek bij het MORE-project (Gravemeijer, Van den Heuvel-Panhuizen & Van der Ploeg, t.p.) laten duidelijk zien dat dit door deskundigen op het gebied van de realistische didactiek ook niet zozeer wordt verwacht.

Ook de uitspraak dat de uitgevers van de drie onderzochte moderne methoden (OR, TT, en GIB) het ontwikkelen van een positieve rekenattitude als doel stellen (Harskamp, 1989, 6), is niet in overeenstemming met hetgeen eerder is gemeld. Alleen GIB en TT hebben gevraagd om een instrument voor de rekenattitude en OR niet (Harskamp, 1988, 43-44). OR blijkt ook maar in beperkte mate belang te hechten aan het betreffende instrument. TT vindt het daarentegen zeer belangrijk en GIB, NAT en NCR vinden het belangrijk (zie Tabel 1). Deze scores hebben echter niet betrekking op het belang van de doelstelling zelf, maar op het belang dat wordt gehecht aan het instrument. Het zegt volgens Harskamp iets over de inhoudsvaliditeit ervan (o.c., 50)[9].

Tabel 1 *Belang toetsen beoordeeld door uitgevers (Harskamp, 1988, 50, tabel 5.4).*

methode	OVS	NAT	NZR	NR	NCR	OR	GIB	TT
CITO-toets Rekenen	-	5	6	3	4	5	4	5
RION-toets Rekenen	-	5	?	3	5	4	5	6

1 = totaal onbelangrijk
2 = niet belangrijk
3 = nauwelijks belangrijk
3 = enigszins belangrijk
5 = belangrijk
6 = zeer belangrijk

	sign. verschil tussen methoden	voor IQ en milieu gecorrigeerde sign. correlaties tussen impl.en lkr.kenm. en toetsprestaties	relevante controle-variabelen
implementatie kenmerken			
% basisstof	+	+	•
lessen per week	-	+	•
minuten per les	-	-	
variatie in leerstof	+	+	
leerstofdifferentiatie	+	-	
leerkracht kenmerken			
ervaring methode	+	-	
leerl.-leerstof attitude	-	+	•
wiskunde in vooropl.	-	-	
leerling kenmerken			
intelligentie	+		•
milieu	-		•

Figuur 4 Overzicht implementatie-, leerkracht- en leerlingkenmerken en gebruikte controlevariabelen.

5.2 *Analyse op klasseniveau*

Bij de analyse op klasseniveau is nagegaan "of door middel van controle voor verschillen in implementatie-, leerkracht- en leerlingkenmerken, die niet het gevolg zijn van de gehanteerde methode, andere resultaten kunnen worden bereikt" dan bij de analyse op leerlingniveau (o.c., 87). Eerst is via de berekening van de partiële correlaties nagegaan welke kenmerken van belang zijn in verband met hun invloed op de klasseprestaties. Een keer is hierbij gecontroleerd voor de gemiddelde intelligentiescore van een klas en een keer voor de gemiddelde milieuscore (o.c., 64). Dit leverde vijf relevante controlevariabelen op: intelligentiescore, milieuscore, aantal rekenlessen, percentage leerlingen met basisstof en leerkrachtattitude (o.c., 65) (zie Figuur 4).

De twee vakdidactische implementatiekenmerken variatie in leerstof en leerstofdifferentiatie zijn niet als controlevariabele opgenomen, omdat deze variabelen mede beïnvloed worden door de factor methode (o.c., 65). Het algemene implementatiekenmerk percentage basisstof, waarvoor ook een significant verschil is gevonden tussen de methoden (o.c., 54), wordt daarentegen wel als controlevariabele meegenomen, terwijl dit kenmerk ook zou kunnen worden gezien als een effect van de methode. In de trant van: een goede methode leidt tot goede leerresultaten bij meer leerlingen, ofwel door een goede methode wordt met meer leerlingen de basisstof behandeld. Harskamp zelf sluit ook niet uit dat er sprake kan zijn van een effect van de methoden, maar beschouwt dit zonder verdere toelichting als een ongewenst, niet door de methoden bedoeld effect (o.c., 85).

De uitgevoerde multivariate covariantie-analyse, waarbij gecontroleerd is voor de genoemde variabelen leverde geen significant methode-effect op ten aanzien van de CITO- en RION-toetsscores van de klassen (o.c., 66).

Hoewel Harskamp het niet erg plausibel acht, zou het volgens hem toch kunnen dat bepaalde methoden enige invloed hebben op het intelligentieniveau van klassen (o.c., 87). Vandaar dat ook nog een vergelijking is uitgevoerd zonder statistische controle voor intelligentie en de andere controlevariabelen. Voor de variabele percentage basisstof is in zekere zin wel gecontroleerd. De analyse is alleen gebaseerd op de klassen waarin met 75% of meer van de leerlingen de basisstof is behandeld. Ook nu weer leverde een multivariate variantie-analyse van de klasseprestaties geen significant verschil op tussen de acht methoden (o.c., 68).

Een secundaire analyse van de gegevens van Harskamp (o.c., 67) wijst echter uit, dat die minimum-eis van 75% bij de traditionele methoden leidt tot gemiddeld 41% teruggang in klassen, terwijl bij de moderne methoden sprake is van een gemiddelde teruggang van 25% (zie Tabel 2).

Tabel 2 *Teruggang aantal klassen per methode en per type methode bij de 75% eis.*

methode	OVS	NAT	NZR	NR	NCR	OR	GIB	TT
totaal aantal klassen	7	12	13	23	19	19	10	12
aantal klassen bij 75% eis	2	8	11	18	7	19	9	3
% vermindering per meth.	71	33	15	22	63	0	10	66
% vermindering per type meth.			41				25	

6 *Invloed implementatiekenmerken op leerresultaten*

Om de laatste onderzoeksvraag – welke verschillen in implementatie van rekenmethoden leiden tot verschillen in leerresultaten? (o.c., 5) – te beantwoorden, zijn alle factoren die samen bleken te hangen met de toetsprestaties van klassen, nadat er voor verschillen in intelligentiescore tussen klassen is gecorrigeerd, in een model gestopt. Dit zijn de implementatiefactoren percentage basisstof, aantal rekenlessen per week, variatie in leerstof en leerstofdifferentiatie (o.c., 71)[10]. Na toetsing van het model met behulp van het programma LISREL IV is de conclusie, dat de rol die de implementatiekenmerken spelen ongeacht de methode, nu duidelijker is dan bij de multivariate covariantie-analyse op klasseniveau. De belangrijkste factor voor de toetsprestaties blijkt het percentage basisstof te zijn. De factoren variatie in leerstof en aantal rekenlessen per week hebben een iets kleiner totaaleffect (o.c., 88).

7 *Aanbevelingen*

Na het onderzoek komt Harskamp met een aantal aanbevelingen. Verder onderzoek is volgens hem van belang om aan te tonen welke concrete lesgedragingen samenhangen met de hoeveelheid behandelde stof en de actieve leertijd van de leerlingen. Deze factoren zouden te zamen een algemeen model voor doelmatig rekenonderwijs vormen, dat een belangrijk gedeelte van de rekenprestaties kan verklaren (o.c., 91). Harskamp verwijst daarbij naar Brophy (1986), terwijl Brophy in plaats van te pleiten voor een algemeen model, juist vindt dat het onderwijsonderzoek minder 'content free' moet worden. Alleen op die manier kan volgens Brophy – die zelf geen reken/wiskunde-didacticus is – het onderzoek het stadium van de kinderschoenen ontgroeien en een relevante basis verschaffen voor de onderwijspraktijk. Wil men verder komen dan moeten vakdidactici en algemeen onderwijskundigen gaan samenwerken, aldus Brophy (1986).

Bij de aanbevelingen richt Harskamp zich ook op de begeleiders en nascholers en op de makers van methoden. "Leerkrachten dienen doelmatig met een methode te kunnen werken in de zin van behandeling van leerstof en besteding van lestijd. Wellicht dat daarna vakdidactische kennis en vaardigheden van belang kunnen zijn om leerresultaten van leerlingen te verbeteren" (Harskamp, 1988, 93). Voor de methoden geldt, dat ze strikter dan nu het geval is, dienen aan te geven welke minimale mate van implementatie van algemene en vakdidactische kenmerken wordt verondersteld voor het bereiken van goede rekenresultaten (Harskamp, 1989, 12)[11]. Harskamp besluit met: "Wanneer methodemakers er in slagen de methode op deze punten te verbeteren, zullen leerkrachten de methode effectiever kunnen gebruiken" (o.c., 12).

Met andere woorden: verbeterde methoden leiden tot een effectiever gebruik van die methoden; een effectiever gebruik van de methoden betekent dat met meer kinderen de basisstof wordt behandeld; vervolgens wordt er weer gecontroleerd voor deze variabele en... worden er weer geen verschillen tussen de methoden gevonden; wel blijkt het percentage basisstof een belangrijke factor te zijn voor de leerresultaten; op grond hiervan kunnen een aantal aanbevelingen worden gedaan gericht op de verbetering van de methoden..., etc.

8 *Tot slot*

Samenvattend kan niet anders worden gezegd, dan dat de in de inleiding vermelde conclusies en aanbevelingen op een zeer zwakke basis zijn gebouwd[12]. Ze wekken ten onrechte de indruk, dat de methoden qua leerstof en qua leerresultaten zo goed als niet verschillen. Over hoe het staat met de implementatie van de rekenmethoden zegt het onderzoek eigenlijk niets. Het heeft eerder betrekking op een aantal algemene onderwijscondities dan op het in praktijk brengen van gewenst onderwijs. De aanbevelingen die worden gedaan zijn deels niet onderzocht en voor zover ze wel voortvloeien uit de onderzoeksresultaten niet erg sterk te noemen.

Opgemerkt moet worden dat bij deze reactie op het onderzoek van Harskamp c.s. de door hen gekozen onderzoeksopzet als uitgangspunt is genomen. De kritiek heeft alleen betrekking op het onderzoek zoals het is uitgevoerd. Zaken als de keuze voor een eenmalige produkt-evaluatie en het ontbreken van gegevens over het feitelijke onderwijsgedrag zijn niet aan de orde gesteld.

Evenzo is buiten beschouwing gelaten, dat ook andere onderzoeken aantonen dat de conclusie dat de methode er niet toe doet niet houdbaar is. Zo blijkt uit de PPON-gegevens dat aan het eind van de basisschool de moderne methode Operatoir Rekenen op 10 van de 27 onderdelen significant hoger scoort dan de traditionele methode Naar Zelfstandig Rekenen (Treffers, 1990; zie ook Treffers, 1988). Ook de uitkomsten van het MORE-onderzoek, die in groep 3 al een methode-effect laten zien (Van den Heuvel-Panhuizen, 1989; Gravemeijer, Van den Heuvel-Panhuizen & Van der Ploeg, t.p.), zijn bij de kritiek niet in het geding gebracht.

Hoe nu verder? Hopelijk hebben de verschillende benaderingen van onderzoek van reken/wiskundeonderwijs in de toekomst elkaar nog wat te bieden en is de hier geuite kritiek op het onderzoek van Harskamp c.s. niet het begin van het einde. Want – zoals Brophy al eerder zei – om verder te komen in het onderzoek van reken/wiskundeonderwijs, is het noodzakelijk dat het tot een samenwerking komt van vakdidactici en algemeen onderwijskundigen. Brophy besluit het in 1986 gepresenteerde AERA-paper dan ook met de woorden: "It is in that spirit that the comments and suggestions made in the present paper have been offered" (Brophy, 1986, 344). Hierbij wil ik me graag aansluiten.

Noten

1. Met de 'moderne methoden' worden hier bedoeld de methoden Taltaal, Getal in Beeld en Operatoir Rekenen. De methoden De Wereld in Getallen, Rekenwerk en Rekenen & Wiskunde die én van recentere datum zijn én meer kenmerken bevatten van de realistische didactiek vallen er dus buiten.
2. Van den Oever (1987, 241) omschrijft bij de beoordeling van het 'Eindrapport Vergelijking van rekenmethoden in het basisonderwijs' de door Harskamp c.s. gevolgde procedure ten aanzien van de methode-inhoud als een Procrustes-situatie.
3. Op een andere plaats wordt een andere operationalisering van de leerstofvariatie in methoden gegeven, nl. de verhouding tussen opgaven over hoofdrekenen en redactierekenen enerzijds en kale sommen voor het inoefenen van algemene rekenvaardigheden anderzijds (Harskamp, 1988, 85).
4. Ook met betrekking tot de door leerkrachten uitgevoerde leerstofvariatie wordt op twee andere plaatsen weer een andere operationalisatie gehanteerd, nl. het percentage van de lestijd dat besteed wordt aan hoofdrekenen en redactieopgaven *versus* (curs., mvdh) het percentage algemene rekenvaardigheden (Harskamp, 1988, 49 en 56).
5. Zonder dat uit toetsingsgegevens blijkt dat dit ook is getoetst, wordt hier later aan toegevoegd dat er bij deze kenmerken ook een verschil bestaat tussen de traditionele en de moderne methoden (Harskamp, 1988, 58-59, 84).
6. Bij de implementatiekenmerken variatie in leerstof en leerstofdifferentiatie (zie noot 5) en bij het leerlingkenmerk intelligentie (volgt nog) gebeurt dit ook.
7. Freudenthal richt zich in deze reactie vooral op de manier waarop in het onderzoek met het kenmerk intelligentie wordt omgesprongen. Twee voorname punten van kritiek zijn: a) de kronkelformulering "Moderne rekenmethoden zullen vaker in klassen worden gebruikt met een wat hoger intelligentieniveau"; Freudenthal vraagt zich af waarom Harskamp voor deze formulering heeft gekozen, terwijl in feite wordt aangetoond dat in klassen waar gewerkt wordt met een moderne methode het intelligentieniveau hoger is; en b) de veronderstelling van Harskamp dat (reken)onderwijs geen invloed heeft op het intelligentieniveau van kinderen en de onvoldoende

onderbouwing van deze veronderstelling. Een meer algemeen punt van kritiek is, dat de impact van moderne rekenmethoden een te ingewikkelde kwestie is om eenvoudigweg met statistiekjes over te beslissen.

8. Op pag. 60 staan andere correlaties: verhoudingen r = .6; meten r = .9.
9. Op pag. 62 lijkt het of Harskamp dit zelf is vergeten. Verwijzende naar tabel 5.4 (o.c., 50) zegt hij, dat het ontwikkelen van een positieve rekenattitude een belangrijke doelstelling is van de meeste methoden.
10. Er wordt hierbij verwezen naar tabel 5.10 (Harskamp, 1988, 64), terwijl in die tabel staat aangegeven dat leerstofdifferentiatie bij correctie voor intelligentie niet significant samenhangt met de toetsprestaties.
11. Uit niets blijkt, dat Harskamp heeft onderzocht of het helpt als dit in de methode wordt aangegeven. Ook is onduidelijk waarop de uitspraak is gebaseerd, dat in vrijwel alle methoden een duidelijke planning ontbreekt van de te behandelen leerstof en de te besteden lessen in het cursusjaar (Harskamp, 1989, 12). Het voorbeeld dat Harskamp (1988, 92) hierbij noemt, dat bij De Wereld in Getallen (WIG) pas in 1986 in een gebruikersbulletin uiteen is gezet wat in deze methode als basisstof en als verrijkingsstof kan worden aangemerkt, klopt in ieder geval niet. Al in de eerste druk (van de deeltjes 1a en 1b) van WIG wordt aangegeven wat tot het minimumprogramma behoort, wat tot het basisprogramma en wat tot het uitbreidingsprogramma.
12. Zie in dit verband ook de SVO-beoordeling van het betreffende RION-onderzoek (Van den Oever, 1987) en de kritische kanttekeningen die twee uitgevers bij dit onderzoek hebben gemaakt (Van Velthoven & Janssen, 1986).

Literatuur

Brophy, J., Teaching and learning mathematics: Where research should be going. *Journal for Research in Mathematics Education*, 1986, *17*, 5, 323-346.

Detterman, D. K. & R. J. Sternberg, *How and how much can intelligence be increased*, Norwood: Ablex, 1982 (geciteerd door Harskamp, 1988).

Gravemeijer, K. P. E., M. van den Heuvel-Panhuizen & D. A. van der Ploeg, *Ander rekenboek, andere rekenprestaties. Een vergelijking tussen twee reken-wiskundemethoden in groep 3*. Utrecht: OW & OC, ter perse.

Harskamp, E. G., *Rekenmethoden op de proef gesteld* (Dissertatie). Groningen: RION, 1988.

Harskamp, E., *Implementatie- en effectonderzoek van nieuwe reken-wiskundemethoden in het basisonderwijs*. Paper t.b.v. Symposium 'Evaluatie en de ontwikkeling van wiskundeonderwijs', Groningen, 23 juni 1989.

Harskamp, E. & C. Suhre, Rekenmethoden op de proef gesteld. *School*, 1987, nov., 20-23.

Harskamp, E. & C. Suhre, Rekenmethoden vergeleken: een effectonderzoek aan het einde van de basisschool. *Pedagogische Studiën*, 1988, *65*, 5, 208-218.

Heuvel-Panhuizen, M. van den & F. Goffree, *Zo rekent Nederland*. Enschede: SLO, 1986.

Heuvel-Panhuizen, M. van den, De eerste uitkomsten. In: E. de Moor (Ed.), *Panama Cursusboek 7*. Utrecht: HMN(SOL)/OW&OC, 1989.

Jong, R. de, *Wiskobas in methoden* (Dissertatie). Utrecht: OW&OC, 1986.

Fullan, M. & A. Pomfret, Research on curriculum and instruction implementation. *Review of Educational Research*, 1977, *47*, 335-397.

Freudenthal, H., *Weeding and Sowing. Preface to a science of Mathematical Education*. Dordrecht/Boston: Reidel, 1978.

Freudenthal, H., *Reactie naar aanleiding van het symposium 'Evaluatie en ontwikkeling van wiskundeonderwijs' – Groningen 23 juni 1989*. Brief, 17 juni 1989.

Lohman, D. F., *Two implications of cognitive psychology for educational measurement*. Adress to the Division of Learning and Instruction Institute for Educational Research in the Netherlands, Utrecht, January 27, 1989.

Oever, A. C. C. van den, Beoordeling van het eindrapport 'Vergelijking van rekenmethoden in het basisonderwijs' uitgevoerd door het RION. *Tijdschrift voor Onderwijsresearch*, 1987, *12*, 4, 239-242.

Slavenburg, J. H., *Onderwijsstimulering en Gezinsactivering*. 's-Gravenhage: SVO, 1986 (geciteerd door Harskamp, 1988).

Streefland, L., *Makro-strukturele verkenningen voor het wiskundeonderwijs*. Utrecht: IOWO, 1980.

Suhre, C., Effectief gebruik van Rekenmethoden: Evaluatieonderzoek in het laatste jaar van de basisschool. *Tijdschrift voor Onderwijsresearch*, 1987, *12*, 6, 344-354.

Suhre, C., Rekenmethoden onderzocht. *Basisschoolmanagement*, 1988. sept., 4-6.

Treffers, A., Cijferen in het rekenonderwijs, toen en nu. *Pedagogische Studiën*, 1982, *59*, 3, 97-115.

Treffers, A. & F. Goffree, Rational analysis of realistic mathematics education – the Wiskobas Program. In: L. Streefland (Ed.), *Proceedings of the ninth international conference for the psychology of mathematics education. Volume II*. Utrecht: OW&OC, 1985, 97-121.

Treffers, A. & E. Feijs, Realistisch rekenwiskundeonderwijs. *Willem Bartjens*, 1987, *6*, 4 (juli), 195-193.
Treffers, A., Over de merkbare invloed van onderwijsmethoden op leerprestaties. In: J. Wijnstra, *Balans van het rekenonderwijs in de basisschool*. Arnhem: CITO, 1988.
Treffers, A., Over het nieuwe programma voor rekenen-wiskunde op de basisschool. *Nieuwe Wiskrant*, 1989, *8*, 3, 10-18.
Treffers, A., *Het voorkomen van ongecijferdheid op de basisschool* (Oratie). Utrecht: VOU/OW&OC, 1990.
Velthoven, J.A.M. van & G.M. Janssen, Het RION-onderzoek: Een valse start of een goed begin? In: E.G.rHarskamp & C.J.M. Suhre, *Vergelijking van rekenmethoden in het basisonderwijs. Eindrapport*. Groningen: RION, 1986.
Walker, D.F. & J. Schaffarzick, Comparing curricula. *Review of Educational Research*, 1974, *44*, 1, 83-111.

Curriculum vitae

M. van den Heuvel-Panhuizen (1950) was werkzaam in het basisonderwijs en in het speciaal onderwijs. Studeerde onderwijskunde aan de Rijksuniversiteit te Utrecht en specialiseerde zich op het gebied van de reken-wiskundedidactiek van de basisschool. Heeft meegewerkt aan het NOT-programma 'Pluspunt', de methode 'Zo reken ik ook!' en SLO-publikaties over rekenen-wiskunde. Thans is zij werkzaam bij de vakgroep OW & OC van de RUU en houdt zich daar onder meer bezig met onderzoek naar gebruik en effecten van rekenwiskundemethoden (MORE-project).

Adres: Rijksuniversiteit Utrecht, Vakgroep OW & OC, Tiberdreef 4, 3561 GG Utrecht.

Manuscript aanvaard 25-1-'90

Summary

Heuvel-Panhuizen, M. van den. 'Research on mathematics textbooks tested.' *Pedagogische Studiën*, 1990, *67*, 271-283

Do the learning results depend on the textbooks as used? This is an actual and also very important question. Harskamp and Suhre's answer to this question is not an affirmative one. They compared eight series of Dutch mathematics textbooks. The result of their research was that there is hardly any difference neither between the content of the textbooks nor between the testscores of the children taught with different textbooks. In the present article a critical reflection has been carried out with respect to the conclusions and recommendations of this research. The executed analysis shows that the conclusions and recommendations of Harskamp and Suhre have been founded too weakly to stand the criticism as given.

Discussie

Kritiek op methodenonderzoek tussen wal en schip

E. HARSKAMP en C. SUHRE
RION, Instituut voor Onderwijsonderzoek
Rijksuniversiteit Groningen

Samenvatting

In haar kritiek op het onderzoek naar reken/wiskundemethoden probeert Van den Heuvel-Panhuizen grond te verschaffen aan de hypothese dat verschillen tussen rekenmethoden ook daadwerkelijk in de rekenresultaten van leerlingen zijn waar te nemen. Als reactie op deze kritiek wordt er op gewezen dat Van den Heuvel-Panhuizen voorbijgaat aan het doel van het onderzoek, namelijk een vergelijking te maken van de effecten van rekenmethoden, zoals ze in de schoolpraktijk worden gebruikt. In onze repliek wordt uiteengezet dat de specifieke analyses van leerinhouden en de metingen van specifieke implementatiekenmerken van moderne methoden die zij in het onderzoek mist, niet aansluiten op de doelstelling van het onderzoek. Aan de conclusies van het onderzoek, dat er slechts op één van de zes afgenomen rekentoetsen verschillen zijn tussen methoden, doet de kritiek geen afbreuk.

Inleiding

In de jaren tachtig zijn er voor het basisonderwijs een groot aantal nieuwe reken/wiskundemethoden op de markt gebracht. De nieuwe methoden verschillen in vakinhoudelijk opzicht van de bestaande rekenmethoden. Wat het leerstofonderdeel Rekenen betreft geven de nieuwe methoden niet zozeer andere leerstof, maar wordt de leerstof op een geheel andere manier in de methoden aangeboden. Nieuwe methoden maken onder andere veel meer gebruik van betekenisvolle thema's ('contexten' genoemd) en van aanschouwelijke schema's om vanuit informele oplossingswijzen van leerlingen te komen tot standaardoplossingen (zie De Jong, 1986 voor een overzicht van vakinhoudelijke verschillen tussen moderne en traditionele rekenmethoden). Voor scholen die een nieuwe rekenmethode moeten aanschaffen, is een belangrijke vraag of moderne en traditionele methoden van elkaar verschillen op aspecten waar het in de dagelijkse lespraktijk op aan komt. Belangrijke aspecten zijn: 1) de uitvoerbaarheid van de geplande rekenlessen in de methoden, 2) het plezier dat leerlingen zullen hebben in de rekenlessen en 3) de rekenvaardigheden die leerlingen verwerven.
Wanneer een nieuwe methode op de markt wordt gebracht bestaat over deze drie aspecten geen objectieve informatie.

Reactie op de kritiek

Het onderzoek waarop Van den Heuvel-Panhuizen (1990) commentaar levert heeft als doel na te gaan in hoeverre er verschillen zijn tussen volledig uitgebrachte rekenmethoden in de leerresultaten van leerlingen aan het einde van de basisschool. Van den Heuvel-Panhuizen richt zich in haar kritiek vooral op het gegeven dat in het onderzoek zo weinig rekening is gehouden met de specifieke doelstellingen en inhouden van de moderne methoden en met de specifieke wijze waarop deze methoden moeten worden gebruikt. We zullen in deze repliek laten zien dat dit ook niet de bedoeling van het onderzoek is geweest. Het gaat ons er in eerste instantie om te laten zien of de verschillende methoden uitvoerbaar zijn in gewone schoolsituaties en in hoeverre verschillende rekenmethoden – zoals ze in de gewone schoolsituaties worden gebruikt – tot verschillen in leerresultaten van groepen leerlingen leiden. Omdat Van den Heuvel-Panhuizen aan deze doelstelling van het onderzoek voorbijgaat, valt haar kritiek tussen wal en schip. Toch is de kritiek interessant met het oog op toekomstig vergelijkend onderzoek van rekenmethoden. We zullen puntsgewijs de gemaakte opmerkingen bespreken.

Het *eerste onderwerp* van kritiek betreft

Pedagogische Studiën 1990 (67) 284-290

onze analyse van de inhoud van de reken/wiskundemethoden. De kritiek richt zich op de kwantitatieve vergelijking van leerstofcategorieën tussen de methoden met voorbijgaan aan een inhoudelijke analyse van de leerstof. Van den Heuvel-Panhuizen doet het daarbij voorkomen alsof de door ons gemaakte leerstofanalyse bedoeld is om voor scholen een keuze uit verschillende rekenmethoden te vergemakkelijken (zie Van den Heuvel-Panhuizen, 1990, p. 273). Deze bedoeling hebben wij met de leerstofanalyse niet gehad. Harskamp (1988, p. 10) zegt in dit verband "Doel is om na te gaan in hoeverre er verschillen zijn tussen methoden in hoeveelheid leerstof op bepaalde rekenonderdelen. Verschillen in nadruk op bepaalde rekenonderdelen geven een indicatie van het belang dat rekenmethoden hechten aan bepaalde leerdoelgebieden". Naar aanleiding van deze analyse is aan uitgevers van de betrokken rekenmethoden een lijst toegestuurd van doelstellingsgebieden. De meeste uitgevers hebben het antwoordformulier in overleg met hun auteurs ingevuld. In overleg met de uitgevers zijn rekentoetsen gekozen en samengesteld voor afname in de scholen. De door ons onderscheiden leerstofcategorieën zijn gebaseerd op a. het CITO doelenboek voor de eindtoets basisonderwijs en b. een publikatie van de Nederlandse Vereniging voor Onderzoek van het Reken Wiskundeonderwijs (Harskamp, 1988 p. 10). Een uitgebreide beschrijving van de leerstofanalyse is gegeven in Wiersma en Willemsen (1986). De analyseresultaten gaven geen aanleiding voor verdere verfijning van leerstofcategorieën. Het gevaar dat Van den Heuvel-Panhuizen signaleert dat er geheel verschillende leerinhouden onder dezelfde leerstofcategorie worden begrepen is voor de leerstof van groep 7 en 8 van de basisschool niet aanwezig. Alle betrokken methodemakers erkennen dat in de bovenbouw sterk rekening is gehouden met de inhoud van de CITO eindtoets (Rekenen en Informatieverwerking) en in de methoden zelf is er qua leerstofonderwerpen grote overeenkomst. Harskamp (1988, p. 38) geeft dit als volgt weer: "De leerstofanalyse laat zien dat er tussen de methoden over het geheel genomen geen grote verschillen zijn in aandacht voor de leerstofcategorieën. Vrijwel alle methoden leggen veel nadruk op Bewerking, Metriekstelsel, Procenten, Hoofdrekenen en Getallen. Er zijn enige verschillen in nadruk tussen methoden op het gebied van Hoofdrekenen, Meten en Schatten van maateenheden, Combinatoriek en Meetkunde. Wat betreft inhoudelijke verschillen in aanbieding van leerstof heeft ons onderzoek zich geheel gebaseerd op de vakinhoudelijke analyse van De Jong (1986). Deze concludeert dat er zeer duidelijke verschillen in Wiskobaskenmerken zijn tussen traditionele rekenmethoden aan de ene kant en moderne rekenmethoden aan de andere kant. Dat is ook de reden geweest waarom we in ons onderzoek deze groepen van methoden hebben vergeleken. Onze kwantitatieve analyse van leerstof laat zien dat de verschillen in aandacht voor de hierboven genoemde leerstofcategorieën tussen de groep traditionele en de groep moderne methoden vaak even groot zijn als binnen deze beide groepen. Van den Heuvel-Panhuizen (1990, p. 273) laat de meest treffende verschillen binnen de groep moderne rekenmethoden (Operatoir Rekenen versus Taltaal) zien. De verschillen binnen de groep moderne methoden geven aan dat men, op grond van verschillen in nadruk op bepaalde leerstofcategorieën, geen hoge verwachting kan hebben ten aanzien van verschillen in toetsprestaties tussen moderne en traditionele methoden. Het is opvallend dat Van den Heuvel-Panhuizen zelf deze voor de hand liggende conclusie niet trekt. Zij doet het voorkomen alsof door verdere specificering van leerstofonderdelen de kwalitatieve leerstofverschillen tussen moderne en traditionele methoden duidelijker zullen worden. De analyses van Wiersma en Willemsen (1986) wijzen echter niet in deze richting. Van den Heuvel-Panhuizen (1990, p. 274) meent verder dat de kwalitatieve beoordelingen van Wiskobaskenmerken in methoden door De Jong (1986) en onze kwantitatieve leerstofanalyse elkaar weerspreken. Ze vergist zich daar. De Jong onderzocht de aanwezigheid en getrouwheid van Wiskobaskenmerken in rekenmethoden voor alle leerjaren van de basisschool. Hij onderscheidt op het gebied van Rekenen, Meten en Meetkunde, Relaties en Functies, Statistiek en Taal en Logica 21 kwalitatieve kenmerken. Harskamp (1988) geeft een analyse van de hoeveelheid rekentaken verdeeld over 14 hoofdcategorieën van leerstof in de methodeboekjes voor groep 7 een 8. Het gaat hier dus om twee geheel verschillende manieren van analyseren van

de inhoud van rekenmethoden[1].

Wanneer Van den Heuvel-Panhuizen (1990, p. 274) kritiek levert op de wijze waarop door ons variatie in leerstof is vastgesteld blijkt dat ze enige moeite heeft met het begrip 100%. In ons onderzoek betekent 100%: het geheel van rekentaken per methode. Variatie in leerstof is afgemeten aan het percentage rekentaken per methode dat wordt besteed aan hoofdrekenen en eigenschapsrekenen en het percentage rekentaken besteed aan redactie opgaven in verhouding tot de overige rekentaken. De overige rekentaken zijn dan rekentaken van 12 leerstofcategorieën (exclusief hoofdrekenen en eigenschapsrekenen) die niet in de vorm van redactieopgaven zijn uitgewerkt. Er is dus geen sprake van dat in onze operationalisering van variatie in leerstof bepaalde leerstofonderdelen in de rekenmethoden buiten beschouwing zijn gebleven. In onze operationalisering van variatie in leerstof in de vragenlijst voor leerkrachten is de omschrijving van 'eigenschapsrekenen' vervangen door 'handig rekenen' met als voorbeeld $118 \times 20 = 120 \times 20 - 2 \times 20$ (zie Harskamp, 1988, p. 127). Het onderscheid tussen hoofdrekenen en handig rekenen is voor vele leerkrachten in de praktijk moeilijk te maken, vooral ook omdat in rekenmethoden deze leerstofcategorieën vaak in dezelfde taken worden aangeboden. Vandaar dat deze categorieën in de leerkrachtenlijst zijn samengetrokken tot één categorie. Onze operationalisering van variatie in leerstof zoals gevraagd aan leerkrachten: percentage *lestijd* besteed aan redactie opgaven en hoofdrekenen (inclusief handig rekenen) in vergelijking tot het percentage lestijd besteed aan oefening van algemene rekenvaardigheden (op het gebied van cijferen, metriekstelsel, meten, breuken, procenten etc.) wijkt dus alleen af van die bij de analyse van leerstof waar het de *maateenheid* betreft. De vraag aan leerkrachten gaat over percentages van de lestijd en de leerstofanalyse gaat over percentages van de leerstof. Vervolgens doet Van den Heuvel-Panhuizen het voorstel om bij de bepaling van variatie in leerstof uit te gaan van slechts 60% van de leerstof of van slechts 80% van de leerstof. Wij vinden op dit punt de kritiek onjuist. Het lijkt ons bepaald niet gerechtvaardigd om variatie in leerstof slechts tot een arbitrair deel van de leertaken uit een methode te beperken. Uitkomsten van analyses zijn op deze manier gemakkelijk te manipuleren.

Het *tweede onderwerp* van kritiek betreft de wijze waarop het gebruik van moderne methoden is vastgesteld. Van den Heuvel-Panhuizen meent dat in het onderzoek de specifieke vernieuwende kenmerken van moderne rekenmethoden onderzocht hadden moeten worden. Over de implementatie van specifieke kenmerken van traditionele methoden rept zij met geen woord. Zij meent (p. 275) dat kenmerken als "het gebruik van contexten", "de aandacht voor modellen", "de grote mate van eigen inbreng van leerlingen", "het interactieve karakter" en "de onderlinge samenhang van leerstofonderdelen" onderzocht moeten worden. Om verschillende redenen hebben wij in ons onderzoek hiervan afgezien.

Ten *eerste* omdat het hier om 'theoretische' begrippen van een goed gebruik van moderne methoden gaat die niet expliciet zijn uitgewerkt in de methoden. Wanneer wordt bijvoorbeeld een 'context' gebruikt volgens de bedoelingen? De door ons onderzochte moderne rekenmethoden werken dit niet expliciet in de handleiding uit. Contexten, modellen en onderlinge samenhang van leerstofonderdelen zijn vooral vernieuwende kenmerken van de *leerstof* in de moderne methoden (zie De Jong, 1986). Methodemakers laten leerkrachten (waarschijnlijk terecht) vrij in de wijze waarop ze de leerstof overdragen. Handleidingen van de moderne rekenmethoden zijn bedoeld als lessuggesties en geven soms alternatieve werkwijzen aan.

In de *tweede* plaats ontbreken empirische gegevens over de wijze waarop de door Van den Heuvel-Panhuizen genoemde theoretische gebruikskenmerken toegepast moeten worden om van belang te zijn voor het verwerven van rekenvaardigheden. Deze theoretische kenmerken zijn tot nu toe nog niet eenduidig geoperationaliseerd en beproefd op hun effectiviteit.

In de *derde* plaats is het onzes inziens niet nodig om bij het vergelijken van de leerresultaten van leerlingen van verschillende methoden aandacht te besteden aan specifieke vernieuwende gebruikskenmerken. Het onderzoek richtte zich immers op vergelijking van methoden zoals leerkrachten ze in normale lessituaties gebruiken. Het is niet de bedoeling van het onderzoek om na te gaan wat de verschillen tussen moderne en traditionele methoden

zijn in scholen waar de (nog te operationaliseren) specifieke vernieuwende didactische kenmerken van de methoden optimaal zijn geïmplementeerd. Van den Heuvel-Panhuizen laat overigens in het midden welke rol deze vernieuwende kenmerken moeten spelen in de analyse van effecten van methoden. Ons lijken deze kenmerken beslist ongeschikt als controlevariabelen. Ze kunnen het best als onlosmakelijk onderdeel van een methode worden gezien en zullen derhalve samenvallen met de factor Methode. Haar kritiek noopt dus niet tot een andere analyse-opzet dan de onze.

In de *vierde* plaats dienen methoden, als ze op de markt gebracht worden, wat hun effect betreft, bestand te zijn tegen verschillen in gebruik van specifieke vernieuwende kenmerken. Wanneer een methode in bepaalde lessituaties (bijvoorbeeld bij klassikaal frontaal onderwijs of bij onderwijs in verschillende niveaugroepen) niet effectief kan werken, dan dienen de scholen hiervoor gewaarschuwd te worden in de vorm van expliciete handleidingen voor goed gebruik van de methode (vergelijkbaar met de bijsluiter bij medicijnen).

Het *derde onderwerp* van kritiek vormt het onderzoek naar de rekenvaardigheid en de rekenattitude van leerlingen die met de verschillende rekenmethoden les hebben gehad. Doel van dit onderzoek was na te gaan of de rekenvaardigheid en de rekenattitude van leerlingen afhangt van de methode waarmee ze les krijgen. Voor de evaluatie van de rekenvaardigheid zijn één algemene rekenvaardigheidstoets en vijf specifieke rekenvaardigheidstoetsen bij de leerlingen afgenomen. In het onderzoek is er vanuit gegaan dat bij de vergelijking van rekenprestaties niet volstaan kan worden met een simpele vergelijking van de scores van leerlingen van de verschillende rekenmethoden aan het einde van de basisschool. Er moet rekening gehouden worden met de mogelijkheid dat verschillen in de toetsprestaties van leerlingen tussen de methoden veroorzaakt kunnen worden door andere oorzaken dan de rekenmethoden. Zo kan miskenning van het belang van verschillen in sociaal milieu en intelligentie tussen de groepen leerlingen die met de verschillende methoden les hebben gehad tot incorrecte conclusies leiden ten aanzien van de vraag welke methode tot betere rekenresultaten leidt. Een voorbeeld kan dit verduidelijken. Wanneer blijkt dat op scholen waar elke week drie uur zwemles wordt gegeven de rekenprestaties beter zijn dan op scholen zonder zwemles, dan zal men de oorzaak voor de betere rekenprestaties eerder zoeken in de verschillen tussen deze twee groepen scholen in leerlingkenmerken en algemene onderwijskenmerken dan in de aanwezigheid van zwemlessen. Wil men het aannemelijk maken dat zwemlessen de oorzaak van de betere rekenprestaties zijn, dan moet men eerst een aantal voor de hand liggende verklaringen ontzenuwen.

Ook bij het aantonen van effecten van rekenmethoden moet men eerst verschillen in leerling- en onderwijskenmerken als alternatieve verklaringen uischakelen en niet net doen alsof de gebruikte rekenmethode de enige aanleiding is voor verschillen in rekenprestaties tussen scholen. In onze analyses van de onderzoeksgegevens zijn eerst leerlingkenmerken en daarna algemene kenmerken van goed gebruik van methoden (zoals het relatieve aantal leerlingen waarmee de basisstof uit de methode is gedaan) als controlevariabelen bij de vergelijking van de rekenprestaties tussen de methoden opgenomen. Er zijn twee soorten analyses uitgevoerd: covariantieanalyses voor de totale groep onderzochte scholen en variantie-analyse voor een homogene groep scholen waar vrijwel met alle leerlingen de stof uit de methode is behandeld[2]. Ten aanzien van het onderzoek naar de effecten van rekenmethoden op rekenprestaties gaan we in op drie kritiekpunten.

Veel kritiek heeft Van den Heuvel op het feit dat in het onderzoek rekening is gehouden met de verschillen in intelligentieniveau van de leerlingen. Zij doet het in deze kritiek voorkomen alsof de onderzoekers stellen dat onderwijs geen invloed heeft op de intelligentie van leerlingen. Van den Heuvel-Panhuizen heeft dan de essentie van het betoog niet goed begrepen. Dat intelligentie voor een deel ook door onderwijs beïnvloed kan worden behoeft geen discussie (zie Harskamp, 1988 p. 59). Waar het om gaat is of verschillen die er tussen scholen bestaan in het intelligentieniveau van de leerlingen geheel of gedeeltelijk beïnvloed zijn door het onderwijs met de gebruikte rekenmethode, los van het onderwijs dat leerkrachten leerlingen geven in andere vakken en vaardigheden. Er is ons geen enkel onderzoek bekend waaruit onomwonden blijkt dat de intelligentie van leerlingen door de verschillen in rekenmethoden worden beïn-

vloed. In ons onderzoek is de intelligentie gemeten met de ISI-test (Boxtel, Snijders & Welten, 1982). De ISI-test bestaat uit subtoetsen waarmee onder meer kennis van woordsynoniemen, tegenstellingen en soortbegrip van woorden wordt gemeten. Deze subtoetsen uit de intelligentie-toets vertonen geen enkele overeenkomst met de inhoud van de moderne rekenmethoden of met de inhoud van de rekentoetsen.

Ten aanzien van de conclusie dat er geen verschillen zijn tussen de methoden in prestaties op de rekentoetsen Meten en Verhoudingen, maakt Van den Heuvel een ongefundeerde opmerking. Zij suggereert namelijk dat onze conclusie over de samenhang van de relatieve aandacht voor deze leerstofonderdelen en de gemiddelde toetsprestaties op deze onderdelen tegenstrijdig is aan onze conclusie dat er geen verschillen tussen methoden zijn in toetsprestaties voor Meten en Verhoudingen. Onze conclusies zijn echter niet tegenstrijdig. De samenhang geeft namelijk aan dat er op het hoogste aggregatieniveau een trend is in de richting dat groepen leerlingen van methoden die relatief meer aandacht aan de twee onderwerpen besteden ook beter presteren ($n = 8$ methoden), terwijl de afwezigheid van significante verschillen tussen methoden in de toetsprestaties op klasniveau ($n = 118$ klassen) aangeeft dat deze trend in leerlingprestaties te verwaarlozen is (de verschillen tussen de methoden zijn namelijk niet significant). Er is geen sprake van dat hiermee voorbij wordt gegaan aan de conclusie van Walker en Schaffarzick (1974), dat methoden verschillen ten aanzien van verschillen in nadruk die op leerstofonderdelen wordt gelegd. Onze gegevens rechtvaardigen echter de conclusie dat op vijf specifieke onderdelen (Meten, Schatten, Verhoudingen, Combinatoriek, en Tabellen en Grafieken) slechts op één onderdeel verschillen zijn in toetsprestaties. Dit is een onderdeel waarop duidelijke verschillen bestaan in leerstof tussen methoden.

Een ander kritiekpunt richt zich op de analyse die is uitgevoerd ten aanzien van de scholen waar met 75% of meer van de leerlingen de basisstof is behandeld. Van den Heuvel-Panhuizen doet alsof deze analyse is uitgevoerd omdat er enige twijfel zou kunnen bestaan over de vraag of bepaalde methoden meer dan andere effect zouden hebben op de intelligentie van leerlingen. Dat is niet de reden waarom deze analyse is uitgevoerd. De reden ligt (ook verwoord in Harskamp & Suhre, 1988) in de statistisch kunstmatige wijze waarop in de vergelijking van effecten tussen methoden voor controle-variabelen wordt gecontroleerd. Door een groep scholen te kiezen waar met minimaal 75% van de leerlingen de basisstof is behandeld worden twee vliegen in één klap geslagen: 1) de groep scholen is meer homogeen (en dus meer vergelijkbaar op de controlevariabelen dan de hele groep scholen) en 2) de vergelijking concentreert zich op gebruikssituaties waarin aan de minimum voorwaarde voor implementatie voldaan is. Het nadeel is dat scholen die een methode onvoldoende hebben geïmplementeerd uit de onderzoeksgroep verwijderd worden. Dit betekent overigens niet, zoals Van den Heuvel-Panhuizen (1990, p. 280) lijkt te suggereren, dat bij de methoden waar veel scholen uitvallen nu alleen de scholen met de intelligentere leerlingen overblijven. Voor de verschillen in leerlingkenmerken tussen de methoden wordt in deze analyse niet gecorrigeerd. Het doel van deze analyse is om lezers, die twijfelen aan de uitkomsten van eerdere analyses en die geen inzicht hebben in de statistische controle die door middel van covariantie-analyse bewerkstelligd wordt, ook langs andere weg resultaten aan te bieden die stroken met de resultaten uit de covariantie-analyses.

Ter afsluiting

De conclusies dat het voor de rekenprestaties en het rekenplezier nauwelijks uitmaakt met welke rekenmethode een leerling les krijgt blijven recht overeind staan. De kritiek van Van den Heuvel-Panhuizen gaat aan de hoofddoelstelling van het onderzoek voorbij en is voor het overige niet empirisch gefundeerd. Zo hoeftrhaar laatste opmerking, dat in andere onderzoeken dan ons onderzoek wel sprake is van verschillen tussen methoden in rekenprestaties van leerlingen niet tegen ons onderzoek te spreken. Wat opvalt aan de door Van den Heuvel-Panhuizen genoemde onderzoeken naar verschillen in rekenprestaties is dat daarin geen rekening gehouden wordt met de klassefactor (de invloed van klassespecifieke leerervaringen)[3] en/of met verschillen in leerlingkenmerken tussen de methoden. Dit zijn ernstige omissies, die de zeggingskracht

van deze onderzoeken aantasten.

Overigens moet uit onze conclusie dat leerlingen met een moderne methode niet beter leren rekenen niet afgeleid worden dat verbetering van het rekenonderwijs door middel van moderne methoden ons inziens onmogelijk is. Nieuwe vakdidactische inzichten moeten echter met algemeen onderwijskundige kenmerken van effectieve instructie in harmonie gebracht worden. Nieuwe methoden dienen daarom eerst grondig te worden beproefd voor ze op de markt komen. Het is een belangrijke taak van methode-onderzoek aanwijzingen te vinden om het rekenen van leerlingen te verbeteren en om aan rekenen/wiskunde nieuwe inhoud te geven. Met Van den Heuvel-Panhuizen zijn we het eens dat daartoe een inspanning nodig is waarbij onderwijskundigen en vakdidactici de handen ineenslaan. Een eerste aanzet om tot kenmerken voor effectief lesmateriaal in reken/wiskundemethoden te komen is gegeven in het proefschrift van Harskamp (1988). Wellicht dat in samenwerking met vakdidactici deze kenmerken nader uitgewerkt en beproefd kunnen worden.

Noten

1. De Jong (1986, p. 238) geeft bijvoorbeeld aan dat vermenigvuldigen en delen niet zijn opgenomen in de traditionele methode Naar Zelfstandig Rekenen (NZR). Daarmee wordt niet bedoeld dat deze leerstof niet in NZR is opgenomen, maar dat de uitwerking volgens Wiskobas-ideeën hier geheel ontbreekt.
2. Lezers die geïnteresseerd zijn in de keuze van geschikte statistische analyseprocedures wordt verwezen naar onder meer Cronbach (1983) en Huitema (1980).
3. In veel onderzoeken waarin methoden worden vergeleken, wordt geen rekening gehouden met het feit dat leerlingen binnen eenzelfde klas gemeenschappelijke leerervaringen hebben waarin ze zich onderscheiden van leerlingen uit andere klassen. Men kan stellen dat de unieke klas waartoe een leerling behoort een 'treatment' vormt die los van de methode van invloed is op de toetsprestaties van leerlingen. Harskamp en Suhre (1986, p. 50) laten zien hoe sterk deze invloed is op de rekenprestaties. Voor een afgewogen oordeel over methode effecten is het belangrijk dat, naast verschillen in leerlingkenmerken en algemene onderwijskenmerken, rekening wordt gehouden met de effecten van de factor 'klas'. Er zijn verschillende manieren waarop voor deze factor gecontroleerd kan worden (zie bijv. Poyner, 1975; Kirk, 1982, hoofdstuk 10 of Goldstein, 1988).

Literatuur

Boxtel, H. W. J., Snijders, Th. V. J., Welten, *ISI-handleiding voor de gehele testreeks en verantwoording van het prestatiegedeelte. Publicatievorm III.* Groningen: Wolters-Noordhoff, 1982.

Cronbach, L. J., *Designing evaluations of educational and social programs.* San Francisco: Jossey-Bass, 1983.

Goldstein, H., *Multilevel models in Educational and Social Research.* London: Charles Griffin & Company Ltd., 1988.

Harskamp, E., *Rekenmethoden op de proef gesteld.* Dissertatie. Groningen: Rijksuniversiteit Groningen, 1988.

Harskamp, E., C. Suhre, *Vergelijking van Rekenmethoden in het basisonderwijs. Eindrapport.* Groningen: Rion, 1986.

Harskamp, E., C. Suhre, Rekenmethoden vergeleken: een effectonderzoek aan het einde van de basisschool. *Pedagogische Studiën,* 1988, 65, 208-218.

Heuvel-Panhuizen, M. van den, Rekenmethoden-onderzoek op de proef gesteld. *Pedagogische Studiën,* 1990, 67.

Huitema, B. E., *The analysis of covariance and alternatives.* New York: Wiley & Sons, 1980.

Jong, R. de, *Wiskobas in methoden* (Dissertatie). Utrecht: Rijksuniversiteit Utrecht, 1986.

Kirk, R. E., *Experimental Design: Procedures for the behavioral Sciences.* Monterey (California); Brooks/Cole Publishing Company. 1982.

Poyner, H. Selecting units of analysis. In: G. B. Borich (Ed.): *Evaluating educational programs and products* (pp. 399-414). Englewood Cliffs, NJ: Prentice Hal, 1974.

Walker, D. F., J. Schaffarzick, Comparing Curricula. *Review of Educational Research,* 1974, Vol. 44, *1*, 83-111.

Wiersma, Y., T. Willemsen, *Vergelijking van rekenmethoden in het Basisonderwijs. Interimrapport 3: Beschrijving en Analyse van acht rekenmethoden.* Groningen: Rion, 1986.

Curriculum vitae

E. Harskamp studeerde na enkele jaren werkzaam te zijn geweest in het onderwijs, onderwijskunde aan de Rijksuniversiteit Groningen. Vanaf 1980 is hij bij het RION werkzaam als coördinator van de

themagroep 'Programma Evaluatie'. Zijn onderzoeksactiviteiten liggen op het gebied van curriculumontwerp en -evaluatie. Verder is hij als adviseur verbonden aan de Adviesraad voor het Basis en Speciaal Onderwijs.

C. Suhre studeerde experimentele psychologie aan de Rijksuniversiteit Groningen. Hij is sinds 1984 werkzaam in de sectie Onderwijsevaluatie van het RION. Momenteel verricht hij onderzoek naar de cijfergeving bij de eindexamens Engels en wiskunde in het HAVO/VWO. Daarnaast is hij als methodoloog betrokken bij diverse onderzoeksprojecten.

Adres van de auteurs: RION, Postbus 1286, 9701 BG Groningen.

Manuscript aanvaard 12-3-'90

Summary

Harskamp, E. & C. Suhre. 'Criticism of research on arithmetic curricula falls between two stools.' *Pedagogische Studiën*, 1990, *67*, 284-290.

In her criticism of our research, Van den Heuvel-Panhuizen suggests that differences between modern and traditional arithmetic curricula can be observed in the test results of students, in spite of our research findings. In this reaction, we point out that Van den Heuvel-Panhuizen misconceives the main objective of the research, which is comparing the effects of arithmetic curricula as actually used in primary schools. We clarify that the specific analysis of content and implementation of modern arithmetic curricula, which are alledgedly lacking, are not in accordance with the main objective of the research. The research outcomes – only one out of six arithmetic tests shows a significant difference between curricula – remain unchallenged.

Mededelingen

Tweede Internationale Spel-symposium

Van 24 tot en met 26 april 1991 vindt te Amsterdam het tweede internationale spel-symposium plaats met als thema 'Play and Intervention'. Dit symposium wordt georganiseerd door het Postacademisch Onderwijs Sociale Wetenschappen (PAOS), in samenwerking met de Vereniging voor Kinder- en Jeugdpsychiatrie (VKJP) en de International Council for Children's Play (ICCP).

Nadere inlichtingen en instructies/formulieren voor het indienen van abstracts zijn te verkrijgen bij PAOS, Postbus 325, 2300 AH Leiden, tel.: 071-143141.

Juniorenstudiedag Anéla

Op 17 december 1990 wordt door Anéla – de Nederlandse Vereniging voor Toegepaste Taalwetenschap – aan de Faculteit der Letteren van de Rijksuniversiteit Groningen een juniorenstudiedag georganiseerd. Thema's zijn o.m. Onderwijs in vreemde talen, Tweede-taalverwerving en tweedetaalonderwijs, en Taal(onderwijs) en techniek.

Nadere inlichtingen zijn te verkrijgen bij Berna de Boer, Talencentrum/RUG, Postbus 716, 9700 AS Groningen, tel.: 050-635960; Nanette Bienfait, Nederlands Instituut/RUG, Postbus 716, 9700 AS Groningen, tel.: 050-635812; Elize Reinink, Prinsesseweg 13, 9717 BA Groningen, tel.: 050-128288.

Inhoud andere tijdschriften

Comenius
10e jaargang, nr. 1, 1990

In memoriam Martinus Jan Langeveld, door S. Miedema
Pluraliteit, pragmatisme en pedagogiek, door G. Biestra
Heropvoedingsrituelen in een kindertehuis, door J. Dekker
De 'opgevoede persoon' – een seksistisch ideaal? door G. Snik
Homo-emancipatie in gevestigde organisaties, door P. Schedler
Fenomenologie en hermeneutiek. Tussen modernisme en postmodernisme. Verslag van de 'Eighth International Human Science Research Conference', Aarhus, 18-22 augustus 1989, door J. de Mul

Pedagogisch Tijdschrift
15e jaargang, nr. 2, 1990

Medische schoolinspectie, gezondheidsvoorlichting en seksuele opvoeding in België in de 19e en 20ste eeuw, door M. Depaepe
De opkomst van schoolartsen in Nederland tussen 1895 en 1920, door E.F.M. Janssen
Artsen en seksuele opvoeding in Nederland, door H.Q. Röling
Toen kwamen de artsen ... ook bij de moeder, door C. Steverlynck
Aspecten van medicalisering en normalisering bij de opkomst van het medisch-opvoedkundig werk in Nederland in het begin van de twintigste eeuw, door A. van der Wurff
Witte jassen en zwarte toga's, door J.J.H. Dekker

Tijdschrift voor Onderwijsresearch
15e jaargang, nr. 1, 1990

Longitudinaal onderzoek van schoolloopbanen, door L.J.Th. van der Kamp en M.J.M. Voeten
De ontwikkeling in het schoolloopbaanonderzoek: een terugblik op een decennium, door J. Dronkers
Het effect van leerkrachtverwachtingen en streefniveaus op schoolloopbanen in het basisonderwijs, door A.A. van der Hoeven-van Doornum, M.J.M. Voeten en P. Jungbluth
Studie-uitval en vertraging bij zes studierichtingen uit het Wetenschappelijk Onderwijs, door U. de Jong en J. Meijer
Doet godsdienst leren? Schoolloopbanen in de gereformeerde sector van het Nederlands onderwijs, door A.B. Dijkstra
Analyse van bijzondere gebeurtenissen in schoolloopbanen: kansen op zittenblijven, door M.J.M. Voeten en Th. Boland

Tijdschrift voor Onderwijsresearch
15e jaargang, nr. 2, 1990

De doelmatigheid van het algemeen voortgezet onderwijs, door A.G.J. van der Torre en B. Kuhry
The threat of scale drift in item banks, door D.N.M. de Gruijter
Interne structuur en constructvaliditeit van belevingsschalen voor Wiskund en Engels, door H. Kuhlemeier, H. van den Berg en F. Teunisse
Intuïtieve, empirische en theoretische redeneringen, door H. Vos

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 3, 1990

Het inzagerecht in het (voortgezet) speciaal onderwijs, door C.M. van Rijswijk
Werken met handelingsplannen in het speciaal onderwijs, door M.J. Schneider en G.M.P. Loots
Reactie: methodologisch incorrect onderzoek naar anatomisch correcte poppen, door M.H. van IJzendoorn en H. Bus
Repliek op correcte en incorrecte kritische kanttekeningen, door H. Baartman, S. van Hekken en S. Bierens

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 4, 1990

Ontwikkelings- en opvoedingsideeën, door F.S. Robijns, P.M.J. van den Brenk, A.E. Jackson, M. Serra en A. Veeneman
Symbolisch spel voor dove peuters, door Y.M. Hooft-van Santen, J. Hoekman en J. Hellendoorn
Indicatiestelling voor kinderpsycho-therapie in een (semi)residentiële instelling: verheldering van een beslissingsproces, door J.M. Atema en L.M.H. Robbroeckx

Ontvangen boeken

Baartman, H., A. Burgess, C. Rümke e.a., *Incest en hulpverlening.* Acco, Amersfoort/Leuven, 1990, *f* 35,–.

Bisschop, C.B., J. Kruk, D.G. van der Steen (red.), *Onderwijs en verantwoordelijkheid voor de toekomst.* Acco, Amersfoort/Leuven, 1990, *f* 17,50.

Bos, K.Th., B.J. van Kesteren, W.G.R. Stoel & C.J.A.J. Vermeulen, *Schoolverzuim. Maatregelen die scholen (kunnen) nemen.* Swets & Zeitlinger, Lisse, 1990, *f* 21,50.

Brunschot, R.A.M., *Begeleiding van beginnende docenten in het technisch onderwijs. Probleemervaring en probleemaanpak* (dissertatie). Technische Universiteit, Eindhoven, 1990, *f* 30,–.

Corte, E. De, L. Verschaffel, H. Schrooten, R. Indemans & E. Hoedemaekers, *LOGO als springplank voor het verwerven van denkvaardigheden bij zesdeklassers.* Acco, Amersfoort/Leuven, 1990.

Damme, W. van (red.), *Extreem moeilijk te begeleiden jongeren.* Acco, Amersfoort/Leuven, 1990, *f* 30,–.

Dijkstra, S., B.H.A.M. van Hout-Wolters, P.C. van der Sijde (Eds.), *Research on instruction. Design and effects.* Educational Technology Publications, Englewood Cliffs, 1990, $ 39,95.

Droogleever Fortuijn, E., De docerende staat. *Onderwijs en macht in de ontwikkeling van Venezuela* (dissertatie). CESO, Den Haag, 1989, *f* 49,50.

Imelman, J.D., e.a., *Cultuurpedagogiek.* Martinus Nijhof, Leiden, 1990, *f* 18,75.

Ippel, M.J. & J.J. Elshout (red.), *Training van hogere-orde denkprocessen.* Swets & Zeitlinger, Lisse, 1990, *f* 39,90.

Kohnstamm, G.A., J.E. Bates, M.K. Rothbart (Eds.), *Temperament in childhood.* John Wiley & Sons, Chichester, 1989, £ 59.00.

Kool, E., *Werken volgens plan in het speciaal onderwijs* (dissertatie). SDU Drukkerij, Den Haag, 1990.

Lowyck, J. & C.M. Clark (Eds.), *Teacher thinking and professional action.* Universitaire Pers, Leuven, 1989, BFrs. 1200.

Luit, J.E.H. van, A. Meijer, Th.G. Hoogeveen (red.), *Onderwijs aan kinderen die moeilijk leren.* Intro, Nijkerk, 1990, *f* 36,50.

Meijer, W.A.J., *Geestelijke stromingen op de basisschool.* Martinus Nijhof, Leiden, 1990, *f* 24,50.

Olthof, Th., *Blame, anger and aggression in children* (disseratie). Stichting Kinderstudies, Groningen, 1990.

Ruijssenaars, A.J.J.M. & J.H.M. Hamers (red.), *Dyslexie. Ernstige lees- en spellingsproblemen.* Acco, Amersfoort/Leuven, 1989, *f* 38,75.

Schuurs, U.R.S., *Leren schrijven voor lezers. Het effect van drie vormen van probleemgericht schrijfonderwijs op de zinsbouwvaardigheid* (dissertatie). Universiteit Twente, Enschede, 1990.

Treffers, A. & E. de Moor, *Proeve van een nationaal programma voor het reken-wiskunde onderwijs op de basisschool.* Deel 2: Basisvaardigheden en cijferen. Zwijsen, Tilburg, 1990, *f* 39,75.

Vedder, P. (Ed.), *Fundamental studies in educational research.* Swets & Zeitlinger, Lisse, 1990, *f* 44,50.

Verhulst, F.C., J.M. Koot, G.W. Akkerhuis, J.W. Veerman, *Praktische handleiding voor de Child Behavior Checklist.* Van Gorcum, Assen, 1990, *f* 49,50.

Voort, T.H.A. van der, *De invloed van televisiegeweld* (2e herziene druk). Swets & Zeitlinger, Lisse, 1990, *f* 34,50.

Cognitieve effecten van het systematisch leren programmeren in LOGO: een onderwijsexperiment bij zesdeklassers

E. DE CORTE, L. VERSCHAFFEL en H. SCHROOTEN
Onderzoekscentrum voor Onderwijsleerprocessen, K.U. Leuven

Samenvatting

De hoofdbedoeling van deze studie was het ontwikkelen van een krachtige LOGO-leeromgeving gericht op het realiseren van de volgende voorwaarden voor het verwerven en de transfer van denkvaardigheden: 1. een voldoende kennis van de domeinspecifieke basisbegrippen van LOGO; 2. beheersing van de beoogde denkvaardigheden binnen de programmeercontext; en 3. deze denkvaardigheden ook in ten minste één ander inhoudsdomein kunnen toepassen. Een systematisch onderwijsexperiment werd opgezet in drie zesdeklassen volgens de voortoets-natoets opzet met controlegroep. In twee experimentele klassen vond LOGO-instructie plaats, met als voornaamste component het aanleren van een programmeerstrategie. In één van deze klassen werd de leerlingen bovendien uitdrukkelijk geleerd om de vaardigheden uit deze strategie te gebruiken in een ander vakgebied. In de controlegroep vond geen interventie plaats. Op het einde van het schooljaar werd vooreerst nagegaan of de vooropgestelde transfervoorwaarden vervuld waren. We stelden vast dat de eerste twee voorwaarden in beide experimentele groepen gerealiseerd waren; de resultaten met betrekking tot de derde voorwaarde waren minder positief. Verder werden verschillende transfertoetsen afgenomen. De analyse van de resultaten wees uit dat transfer werd bereikt in beide experimentele groepen; dit suggereert dat het vervullen van de eerste twee vermelde voorwaarden voldoende is om transfer van denkvaardigheden te bekomen.

1 *Theoretische achtergrond*

Het verwerven van denk- en probleemoplossingsvaardigheden is een van de kerndoelstellingen van het basisonderwijs. Recent onderzoek heeft uitgewezen dat leren probleemoplossen neerkomt op het geïntegreerd verwerven en leren toepassen van drie soorten vaardigheden (De Corte & Verschaffel, 1987; ter perse; zie ook Perkins & Salomon, 1989), m.n.: 1. het flexibel kunnen hanteren van een goed georganiseerd domeinspecifiek kennisbestand, 2. de beheersing van heuristische methoden, en 3. het beschikken over metacognitieve kennis en vaardigheden. Daarenboven veronderstelt men vaak dat wie deze vaardigheden verworven heeft in een bepaald gebied, ze ook kan toepassen in andere contexten, m.a.w. er wordt verwacht dat *transfer* zal optreden.

De voorbije jaren is veelal beweerd dat leren programmeren een aangewezen hulpmiddel is voor het verwerven en de transfer van belangrijke heuristieken en metacognitieve vaardigheden. Eén van de meest uitgesproken verdedigers van deze 'cognitieve effectenhypothese' met betrekking tot leren programmeren is Papert. In zijn bekende boek *Mindstorms* (1980), waarvan in 1984 een Nederlandstalige vertaling verscheen, verdedigt hij de stelling dat leren programmeren in LOGO leidt tot de ontwikkeling van algemeen bruikbare denk- en probleemoplossingsvaardigheden, zoals plannen en 'debuggen', en tot het verwerven van krachtige begrippen, zoals variabele en recursie. Een ander kernidee uit Paperts visie is dat leren programmeren moet gebaseerd zijn op het principe van het zelfontdekkend leren: kinderen moeten hun kennis zelf actief opbouwen, door alleen en/of samen met andere leerlingen al spelend en experimenterend met de computer om te gaan; systematische sturing van het leerproces is daarbij uit den boze.

De resultaten van de pionierstudies over de cognitieve effecten van leren programmeren die in de eerste helft van de jaren tachtig ondernomen werden, boden evenwel weinig

steun voor deze optimistische hypothese. Een nadere analyse van deze studies suggereert echter dat de teleurstellende resultaten wellicht te wijten waren aan een aantal tekorten van de onderzoekingen zelf. Ten eerste hadden de leerlingen in deze studies onvoldoende programmeerkennis en -vaardigheden verworven, enerzijds door de korte duur van de instructie (gemiddeld ongeveer 25 tot 30 uren) en anderzijds door de afwezigheid van systematische begeleiding van het leerproces. Ten tweede ontbrak een expliciete gerichtheid op transfer (De Corte & Verschaffel, 1987; ter perse).

In 1985 startten we in het Leuvens Onderzoekscentrum voor Onderwijsleerprocessen een project over de cognitieve effecten van leren programmeren in LOGO. De voornaamste doelstelling was het ontwerpen, implementeren en evalueren van een krachtige onderwijsleeromgeving, gericht op het verwerven en de transfer van algemene denkvaardigheden. Daarbij richtten we ons op een subset van de probleemoplossingsvaardigheden waarvan verwacht wordt dat deze beïnvloed worden door het leren programmeren: twee metacognitieve vaardigheden (plannen en debuggen), en twee heuristieken (probleemdecompositie en het maken van een externe probleemrepresentatie).

Op basis van de voorafgaande studies over de effecten van leren programmeren en van de beschikbare literatuur over transfer in het algemeen, stelden we dat voor het bereiken van transfer via LOGO het vervullen van de volgende drie voorwaarden van belang is:

1. de leerlingen hebben voldoende domeinspecifieke kennis verworven (i.e. de primitieven en concepten van de LOGO-taal);
2. ze beheersen de beoogde heuristieken en metacognitieve vaardigheden binnen de context van het programmeren in LOGO;
3. ze hebben uitdrukkelijk geleerd de denkvaardigheden ook in ten minste één ander inhoudsdomein toe te passen.

Omdat uit de voorgaande studies reeds voldoende gebleken was dat het realiseren van de eerste voorwaarde inderdaad een noodzakelijke, doch geen voldoende voorwaarde is voor het bekomen van transfer, hebben we ons toegespitst op het onderzoeken van de bijdrage van de laatste twee condities. De transferhypothese werd meer specifiek als volgt geformuleerd: "Indien aan de eerste twee voorwaarden voldaan is, treedt er transfer van de beoogde denkvaardigheden op; aanvullende realisatie van de derde voorwaarde leidt tot een significant hoger transfereffect".

Het project werd uitgevoerd in twee fasen. Gedurende het schooljaar 1986-87 werd een *vooronderzoek* uitgevoerd, gericht op het ontwikkelen, implementeren en evalueren van een LOGO-onderwijsleeromgeving en op de constructie van toetsen over de beheersing van LOGO-kennis en -vaardigheden (transfervoorwaarden 1 en 2) (zie De Corte, Verschaffel, Hoedemaekers, Indemans & Schrooten, 1988). Om de transferhypothese te toetsen werd in de tweede fase (schooljaar 1987-88) een meer *systematisch onderwijsexperiment* uitgevoerd, waarin de geconstrueerde leeromgeving uitgebreid werd met een cruciale component, nl. het expliciet nastreven van transfer (transfervoorwaarde 3). In onderhavige bijdrage gaan we in op de opzet en de resultaten van deze tweede studie. Voor een meer uitvoerige beschrijving van dit onderzoek verwijzen we naar De Corte, Verschaffel, Schrooten, Indemans en Hoedemaekers (1990).

2 *Proefgroep en onderzoeksopzet*

2.1 *Proefgroep*

Het onderzoek vond plaats in drie zesdeklassen, elk bestaande uit 24 jongens van 11-12 jaar. De leerlingen hadden geen voorafgaande ervaring met computers in het algemeen, noch met LOGO in het bijzonder. Twee klassen fungeerden als experimentele groepen (E1 en E2); een derde klas vormde de controlegroep (C). In het begin van het schooljaar werden een intelligentietest en een schoolvorderingentoets afgenomen in de drie klassen. Er werden geen significante verschillen vastgesteld.

2.2 *Opzet*

Het toetsingsonderzoek werd uitgevoerd in de vorm van een voortoets-natoets opzet met controlegroep. In E1 werden de eerste twee voorwaarden voor transfer nagestreefd in een 60 uur durende LOGO-cursus, waarin de basisprimitieven en -concepten van de programmeertaal werden aangeleerd, evenals een strategie voor het systematisch schrijven van

LOGO-programma's. In het kader van deze programmeerstrategie, bestaande uit een planfase en een geïntegreerde uitvoerings- en controlefase, trainden we de beoogde denkvaardigheden (plannen, debuggen, probleemdecompositie, en constructie van een externe probleemrepresentatie). In E2 werd daarenboven de derde transfervoorwaarde nagestreefd: de jongens uit deze klas leerden de verworven denkvaardigheden ook nog toepassen in een andere context, nl. bij het oplossen van complexe rekenvraagstukken. In de controlegroep vond geen interventie plaats.

3 *LOGO-leeromgeving*

Met het oog op de verwezenlijking van de eerste twee voorwaarden voor het bereiken van transfer, werd in beide experimentele groepen gedurende het ganse schooljaar elke week een LOGO-namiddag georganiseerd. De instructie werd gegeven door het onderzoeksteam in samenwerking met de klasleerkracht. De computerklas, die direct toegankelijk was vanuit het gewone klaslokaal, was uitgerust met negen Philips MSX-computers. De LOGO-cursus bestond uit twee belangrijke componenten, nl. het aanleren van de LOGO-primitieven en -concepten (eerste transfervoorwaarde) en het trainen van een programmeerstrategie (tweede transfervoorwaarde).

3.1 *Aanleren van LOGO-primitieven en -concepten (eerste transfervoorwaarde)*

De inhoud van de LOGO-cursus werd beperkt tot de zogenaamde 'turtle-graphics', of het tekenen met de schildpad. Bovendien werden enkel deze primitieven aangebracht, die we noodzakelijk achtten om de beoogde programmeerstrategie met succes te kunnen aanleren. Meer bepaald kwamen de volgende LOGO-noties aan de orde: maken van eenvoudige tekeningen met de elementaire primitieven (VT/AT, LI/RE, PO/PN), werken met het HERHAAL-commando, schrijven van procedures en superprocedures.

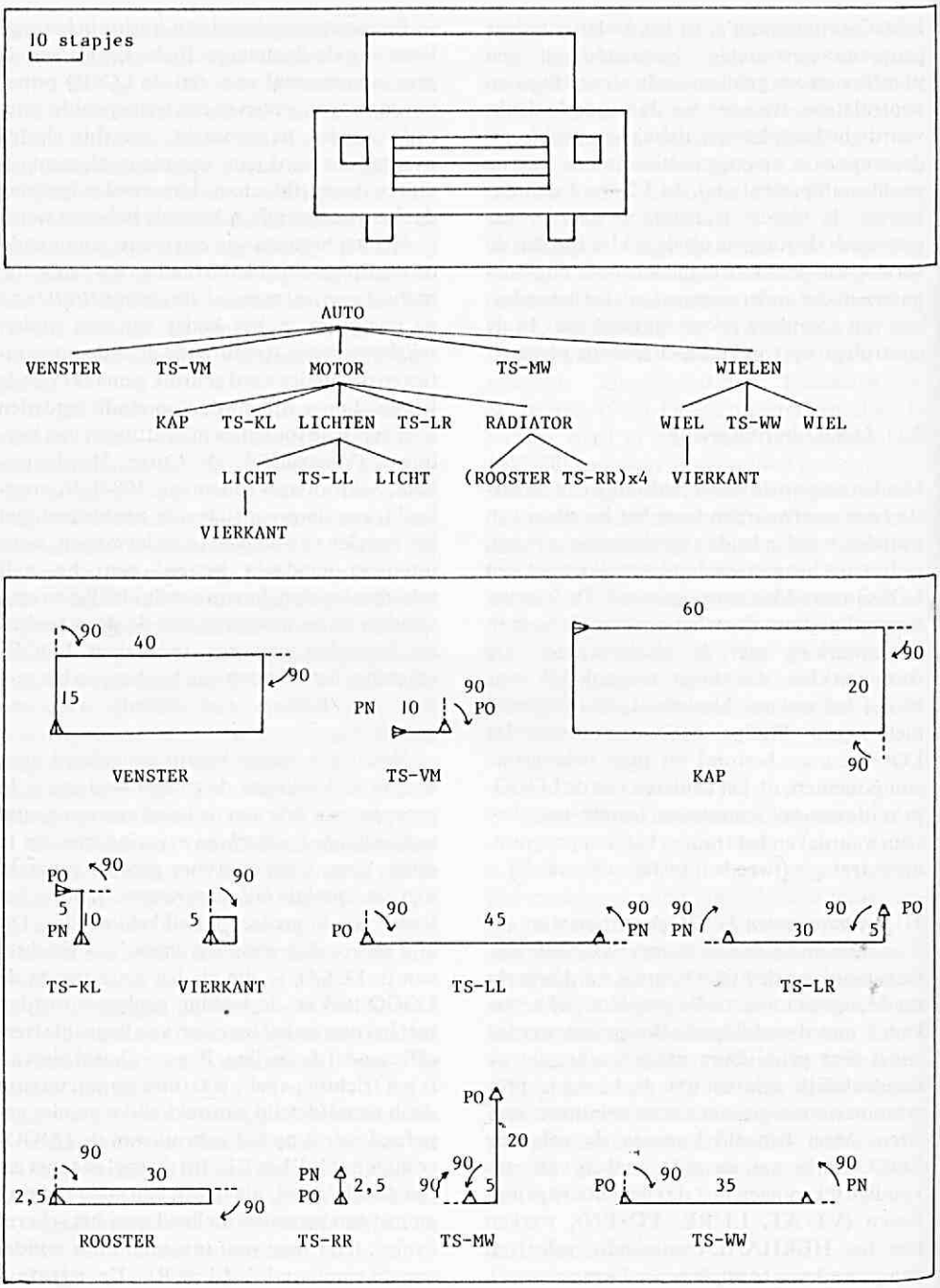
Uitgaande van een gematigd constructivistische visie op leren (De Corte, 1990), gingen we ervan uit dat een krachtige onderwijsleeromgeving gekenmerkt wordt door een evenwicht tussen zelfontdekkend leren enerzijds, en meer systematische instructie en sturing anderzijds.

Een eerste voorbeeld van sturing in het aanleren van de domeinspecifieke kennis van de programmeertaal was, dat de LOGO-primitieven en -concepten in een welbepaalde volgorde werden aangebracht, waarbij slechts overgegaan werd naar een nieuw element als uit een tussentijdse toets kon worden besloten dat het voorgaande voldoende beheerst werd.

Het aanbrengen van een nieuw commando of begrip geschiedde bovendien op een systematische wijze, meestal via *demonstratie* aan de computer in het kader van een *onderwijsleergesprek*. Gedurende deze demonstraties en discussies werd gebruik gemaakt van de kennis die we tijdens de voorstudie opdeden over typische fouten en misvattingen van leerlingen (Verschaffel, De Corte, Hoedemaekers, Schrooten & Indemans, 1987). Bijvoorbeeld, om de geconstateerde problemen met het bepalen van hoeken te ondervangen, werd uitvoerig aandacht besteed aan: het zelf schildpad spelen, het op een duidelijke en eenvoudige wijze aangeven van de draaihoeken op het plan van een te maken LOGO-tekening, het schatten van hoeken, en het meten van hoeken met behulp van een graadboog.

Nadat een nieuw begrip aangeleerd was, kregen de leerlingen de gelegenheid om er in groepjes van drie aan de hand van een aantal *welomlijnde opdrachten* ervaring mee op te doen. Soms werd daarvoor gebruik gemaakt van z.g. speciale *hulpprogramma's*, die in het kader van dit project ontwikkeld werden. Dit zijn eenvoudige spelprogramma's – geschreven in LOGO – die als het ware tussen de LOGO-taal en de leerling geplaatst worden met het oog op het bereiken van bepaalde leereffecten bij de leerling. Een voorbeeld hiervan is het 'richting-spel', waarmee op een intense doch terzelfdertijd aantrekkelijke manier geoefend wordt op het gebruik van de LOGO-commando's LI en RE. Bij dit spel gaat het erom de schildpad, die in een bepaalde richting en met een bepaalde snelheid over het scherm kruipt, naar haar nest te sturen door middel van de commando's LI en RE. De oefenfase werd afgesloten met een leergesprek, waarin ervaringen, moeilijkheden, vondsten, strategieën, enz. van de leerlingen samengebracht en besproken werden.

Met het oog op kennisintegratie kregen de leerlingen telkens de gelegenheid om een nieuw verworven LOGO-begrip samen met



Figuur 1 *Plan voor het schrijven van een LOGO-programma voor de tekening van de auto.*

het eerder geleerde toe te passen in een meer omvangrijk *zelfgekozen* project.

3.2 *Aanleren van een strategie voor het schrijven van LOGO-programma's (tweede transfervoorwaarde)*

De aangeleerde strategie voor het schrijven van een LOGO-programma omvat twee fasen: 1. een planfase, die los van de computer plaatsvindt, en 2. een geïntegreerde uitvoerings- en controlefase aan de computer.

In de *planfase* onderscheiden we de volgende drie stappen (zie Figuur 1):

- Maken van een tekening van wat er precies op het scherm moet verschijnen.
- Opstellen van een boomdiagram waarin de complexe tekening opgedeeld wordt in delen, totdat men louter elementaire bouwstenen bekomt waarvan de programmering in LOGO als 'kinderspel' kan beschouwd worden (bijv. vierkant). Dit diagram duidt meteen de volgorde aan waarin de diverse onderdelen op het scherm getekend moeten worden. In het boomdiagram worden tevens de 'tussenstukken'[1] vermeld die de samenstellende delen verbinden.
- Apart tekenen van elke component, waarbij voor elk deel en elk tussenstuk de grootte van de afstanden en de draaihoeken worden aangegeven, evenals de begin- en eindpositie.

In deze eerste fase komen de volgende denkvaardigheden aan bod: plannen, decompositie (opsplitsen van een complexe figuur in delen) en het maken van een externe representatie (tekenen van een boomdiagram). De vierde vaardigheid die we beoogden te trainen, nl. het debuggen, vinden we terug in de tweede fase.

Pas nadat al de genoemde voorbereidingsactiviteiten beëindigd zijn, gaat de leerling naar de computer om het programma in te tikken. In de *uitvoerings- en controlefase* van de programmeerstrategie worden de volgende twee nauw met elkaar verbonden principes gehanteerd: 1) neerwaarts programmeren en 2) geïntegreerd uitvoeren en controleren.

- *Neerwaarts programmeren.* In plaats van een opwaartse aanpak, waarbij eerst alle enkelvoudige procedures van het laagste niveau van het boomdiagram geschreven worden en pas daarna overgegaan wordt tot het definiëren van de hoger gelegen procedures, hebben wij gekozen voor een neerwaartse aanpak. Dit houdt in dat begonnen wordt met het schrijven van de procedure uit het hoogste niveau van het boomdiagram; dit noemen we de 'moederprocedure'. Daarna wordt achtereenvolgens elke component uit de moederprocedure nader gespecificeerd, totdat het laagste niveau van het boomdiagram bereikt wordt.
- *Regelmatig controleren en debuggen.* In plaats van eerst de moederprocedure en alle deelprocedures in te tikken en pas daarna over te gaan tot het controleren en zonodig verbeteren van alle invoer, hebben wij voor een geïntegreerde uitvoerings- en controleprocedure gekozen. Dit houdt in dat elke nieuwe procedure onmiddellijk getest wordt via het intikken van de naam van de moederprocedure. Het doel daarvan is tweeledig: 1) op het scherm verschijnt het resultaat van de tot dan toe ingevoerde procedures, wat de gelegenheid biedt om te controleren of de tekening en de schildpadpositie tot dan toe correct zijn; en 2) in de foutmelding ("ik ken geen. . .") staat vermeld welke procedure vervolgens geschreven moet worden.

Wat deze principes concreet inhouden, illustreren we weerom aan de hand van de taak uit Figuur 1. Overeenkomstig het principe van neerwaarts programmeren, tikt de leerling eerst de moederprocedure in:

```
LEER AUTO[2]
    VENSTER TS-VM MOTOR TS-MW
    WIELEN EIND
```

Zoals het tweede principe voorschrijft, wordt vervolgens de moederprocedure (AUTO) uitgevoerd. Nadat de leerling het commando AUTO heeft ingetikt, verschijnt op het scherm de mededeling: "ik ken geen VENSTER in AUTO". Daarop definieert de leerling de eerste procedure uit de moederprocedure:

```
LEER VENSTER
    HERHAAL 2 [VT 15 RE 90 VT 40 RE
    90] EIND
```

Wanneer vervolgens weerom de naam van de moederprocedure (AUTO) ingetikt wordt, tekent de schildpad het venster. De foutmelding geeft aan dat de procedure voor de verbinding tussen het venster en de motor (TS-VM) de volgende procedure is die geleerd moet worden. De leerling tikt in:

LEER TS-VM
PO RE 90 AT 10 PN EIND

Bij de daaropvolgende toetsing (AUTO) wordt het venster getekend en gaat de schildpad daarna naar de beginpositie voor de motor. Dan schrijft de leerling de procedure voor de motor. Bijvoorbeeld:

LEER MOTOR
KAP TS-KL LAMPEN TS-LR RADIATOR EIND

Als daarna de moederprocedure (AUTO) wordt uitgevoerd, tekent de schildpad opnieuw het venster en het verbindingstuk tussen het venster en de motor, en geeft vervolgens de boodschap "ik ken geen KAP in MOTOR". De volgende procedure die geleerd moet worden is de procedure KAP.

Dit procédé wordt herhaald tot alle delen en tussenstukken van MOTOR en vervolgens alle bestanddelen van de overige componenten uit de moederprocedure gedefinieerd zijn.

Het expliciet aanleren van deze programmeerstrategie was de belangrijkste component van de LOGO-cursus. Het instructiesysteem werd reeds een eerste keer uitgeprobeerd in het vooronderzoek (De Corte, e.a., 1988). Dit leidde tot de volgende bevindingen. Een groot deel van de leerlingen maakte niet spontaan een boomdiagram wanneer ze voor een complexe figuur een programma moesten schrijven. Dit kwam omdat zij het nut van de aangeleerde strategie niet inzagen. Maar ook wanneer we hen daar expliciet om vroegen, waren verschillende leerlingen niet in staat een correct boomdiagram te maken. Rekening houdend met deze ervaringen en bevindingen, werd de leeromgeving op enkele punten bijgestuurd: in het eigenlijke onderzoek werd de beoogde strategie op een meer systematische wijze aangebracht, gebruik makend van een aantal technieken die in de recente literatuur over het aanleren van denkvaardigheden naar voren komen als zeer belangrijk, zoals 'modelling', 'scaffolding', 'coaching', 'fading', 'articulation', 'reflection' and 'exploration' (Collins, Brown & Newman, in press). Daarnaast trachtten we de leerlingen meer te motiveren om de aangeleerde werkwijze te gebruiken.

Eerst werd de strategie in zijn geheel door een 'expert' (een lid van het onderzoeksteam) gedemonstreerd en werden de verschillende componenten geëxpliciteerd. Daardoor konden de leerlingen een mentaal model van de noodzakelijke stappen en processen ontwikkelen, dat ze bij het zelfstandig toepassen voor ogen konden houden (*'modelling'*).

Daarna werd elke component van de strategie gedetailleerd behandeld en afzonderlijk ingeoefend; zo nodig werd directe hulp geboden bij het uitvoeren van de taak (*'scaffolding'*). Dit hield o.m. in dat de leerkracht in een eerste fase die aspecten van de strategie voor zijn rekening nam die de leerlingen nog niet beheersten. Ook werd in de beginfase van het inoefenen van de planstrategie een computerprogramma gebruikt, dat hulp biedt bij het opstellen van een boomdiagram. Dit programma, dat in het kader van onderhavig onderzoek ontwikkeld werd, stuurt het proces van het opsplitsen van een tekening in delen door gerichte vragen te stellen (bijv. Is MOTOR eenvoudig genoeg? Uit welke delen bestaat MOTOR?). Via het meermaals gebruiken van dit programma voor het maken van boomdiagrammen voor verschillende tekeningen, werd ernaar gestreefd dat de leerlingen de werkwijze voor het plannen zouden internaliseren, om ze daarna zelfstandig te kunnen toepassen.

Vervolgens oefenden de leerlingen in groepjes van twee à drie de volledige programmeerstrategie in aan de hand van steeds moeilijkere tekeningen (*'exploration'*). Daarbij konden zij elk moment teruggrijpen naar kaarten waarop de voornaamste stappen van de strategie vermeld stonden, tesamen met nuttige tips om de onderscheiden fasen goed door te komen. Bij de uitvoerings- en controlefase aan de computer namen de leerlingen afwisselend de rol van 'uitvoerder' en van 'controleur' waar. Aanvankelijk werden de groepjes intensief begeleid (*'coaching'*), o.a. door het geven van aanwijzingen en terugkoppeling; doch afhankelijk van de individuele mogelijkheden en behoeften van de leerlingen verminderde deze hulp van de begeleiders gaandeweg (*'fading'*). Bij het geven van hulp werd expliciet rekening gehouden met de mogelijke fouten, misvattingen en moeilijkheden die uit het vooronderzoek gebleken waren. Zo werden bijvoorbeeld de beoordelingscriteria en de foutenclassificatie voor boomdiagrammen die afgeleid werden uit het vooronderzoek, aan de leerlingen kenbaar gemaakt en frequent gehanteerd – zowel door de begeleiders als door de leerlingen zelf – bij de

bespreking van de boomdiagrammen die gemaakt werden in het kader van de verschillende projecten in de loop van het schooljaar.

Gedurende de hele instructie en training van de strategie werden de leerlingen gestimuleerd om de procedures en technieken die ze aanwendden te verwoorden (*'articulation'*), en om deze te vergelijken met die van andere kinderen en met het expertmodel (*'reflection'*). Een belangrijke instructietechniek met het oog op het bereiken van transfer, was het expliciet en intentioneel *abstraheren* van de getrainde denkvaardigheden.

Om de leerlingen sterker te motiveren tot het gebruiken van de aangeleerde strategie werd ruime aandacht besteed aan het expliciet aantonen van het nut en de waarde ervan. Zo werden oefeningen gemaakt waarbij de leerlingen programma's moesten schrijven voor lichte variaties op eenzelfde tekening, waarbij de voordelen van het planmatig en modulair werken tot uiting kwamen. Ook werd er gedemonstreerd dat het eigenlijke programmeerwerk aan de computer en het verbeteren van fouten veel gemakkelijker en efficiënter verloopt wanneer uitgegaan wordt van een volledig en correct plan.

Om te vermijden dat de leerlingen een aversie zouden ontwikkelen voor een te rigide en strak opgelegde strategie, lieten we toe dat ze de aangeleerde werkwijze minder nauwgezet volgden naarmate het schooljaar vorderde. Zo moesten bijvoorbeeld in het begin alle tussenstukken beginnen met PO en eindigen met PN; naarmate de leerlingen onderscheid leerden maken tussen situaties waar dit noodzakelijk is en waar niet, mocht dit achterwege gelaten worden. Dit is perfect in overeenstemming met onze opvatting over het leerproces. Immers, we beschouwen de volledige procedure als een geschikte aanvangshandelingsstructuur, waarin een adequaat oplossingsproces zeer uitvoerig veruitwendigd wordt. Doch in de loop van het leerproces kunnen allerhande verkortingen en verinnerlijkingen plaatsvinden.

4 *Expliciet nastreven van transfer (derde transfervoorwaarde)*

In de tweede experimentele groep werden op het einde van de LOGO-cursus een vijftal lessen ingelast, waarbij de leerlingen de algemene denkvaardigheden onderwezen in de LOGO-context (nl. plannen, debuggen, probleemdecompositie en constructie van een externe representatie) ook leerden toepassen bij het oplossen van complexe rekenvraagstukken. Meer bepaald leerden ze een strategie bestaande uit twee fasen: een planfase en een uitvoerings- en controlefase. Zoals in LOGO, bestaat het plannen van een oplossingsweg voor een complex vraagstuk erin het probleem op te splitsen in deelproblemen (decompositie-heuristiek); het resultaat van deze planactiviteiten kan eveneens weergegeven worden in een boomdiagram (maken van een externe representatie). In de uitvoerings- en controlefase wordt successief ieder deelprobleem opgelost en worden de antwoorden onmiddellijk gecontroleerd.

5 *Meetinstrumenten*

5.1 *Criteriumtoetsen*

Vooraleer over te gaan tot de toetsing van het transfereffect werd onderzocht of aan de transfervoorwaarden voldaan was. In beide experimentele klassen werden één maand voor het einde van de leergang verschillende criteriumtoetsen afgenomen in verband met de eerste twee transfervoorwaarden. Voor elke toets werd per klas onderzocht of meer dan 75% van de leerlingen het vooropgestelde individueel criterium (nl. 2/3 van het maximum aantal punten) behaalde. Wanneer dit niet zo was, beschouwden we de betreffende voorwaarde als niet voldoende vervuld. In dit laatste geval volgde hierop een remediëring en hertoetsing voor die leerlingen die het criterium niet bereikten. De toetsen in verband met vraagstukken (derde transfervoorwaarde) werden slechts één keer afgenomen, daar er geen tijd beschikbaar was voor remediëring en hertoetsing; dit gebeurde vanzelfsprekend alleen in E2.

De toets over de kennis van de LOGO-primitieven en -concepten (transfervoorwaarde 1) bestond uit twee reeksen van opgaven: bij de eerste reeks moesten de leerlingen aanduiden welk van de vier alternatieve LOGO commando's of procedures een gegeven schermeffect zou produceren; bij de tweede reeks was de opdracht de tekening aan te duiden die door een gegeven programma op het scherm gemaakt zou worden.

Daarnaast waren er drie toetsen om na te gaan of de leerlingen de programmeerstrategie verworven hadden (transfervoorwaarde 2). Twee ervan hadden betrekking op de planfase. In de eerste toets bestond elke opgave uit een tekening met twee boomdiagrammen, waarvan slechts één correct was; gevraagd werd het juiste boomdiagram voor de betreffende tekening aan te duiden. In de tweede toets over de planfase moesten de leerlingen zelfstandig een adequaat boomdiagram construeren voor een gegeven tekening. De derde toets betrof de uitvoerings- en controlefase. In tegenstelling tot de twee voorgaande toetsen, werd deze toets individueel aan de computer afgenomen. De opdracht was een LOGO-programma te schrijven voor een relatief complexe tekening, uitgaande van een gegeven plan. Om het debuggingsproces te kunnen analyseren, werden er in dit plan vier fouten verwerkt.

Om de beheersing van de strategie voor het oplossen van samengestelde rekenvraagstukken (transfervoorwaarde 3) na te gaan, werden analoge toetsen gebruikt. Zoals bij LOGO hadden twee toetsen betrekking op de planfase: in de eerste moesten de leerlingen het beste van twee gegeven boomdiagrammen aanduiden voor een reeks van vraagstukken; in de tweede toets werden ze gevraagd zelfstandig een adequaat boomdiagram te construeren voor een reeks gegeven vraagstukken. De derde toets betrof de uitvoerings- en controlefase; de opdracht was een aantal vraagstukken op te lossen, uitgaande van gegeven boomdiagrammen.

5.2 *Transfertoetsen*

De leerlingen uit de drie proefgroepen kregen een aantal toetsen aangeboden om na te gaan in welke mate ze in staat waren de in E1 en E2 aangeleerde denkvaardigheden spontaan en zelfstandig toe te passen in nieuwe probleemsituaties.

De doolhof-toets, een computertoets voor planvaardigheid, werd individueel aan de machine afgenomen. De leerling kreeg achtereenvolgens drie opgaven aangeboden, waarbij een hongerige kever via de kortste weg naar drie slaatjes gestuurd moest worden. De doolhoven waren zo geconstrueerd, dat plannen noodzakelijk was om een goed resultaat te bekomen.

In de foutendetectie-toets werd zowel het plannen als debuggen gemeten. In drie opgaven werd telkens een dagelijkse situatie geschetst, waarbij persoon X aan persoon Y een reeks instructies gaf om een bepaalde taak te volbrengen (McCoy Carver, 1986). Persoon Y voerde deze instructies nauwgezet uit, maar het verkregen resultaat kwam niet overeen met het gewenste, doordat persoon X een verkeerde instructie gegeven had. De opdracht was om deze foutieve instructie op te sporen en de instructielijst zodanig te veranderen, dat het gewenste resultaat wel bereikt zou worden.

In de blokken-toets bestond de opdracht erin een reeks gegeven complexe figuren te beschrijven in termen van hun samenstellende delen, daarbij gebruik makend van een reeks gegeven elementen. Deze activiteit vereiste de toepassing van de decompositie-heuristiek.

De weekdagen-toets omvatte een reeks opgaven, die alle een variant waren op de volgende vraag: "Welke dag was het eergisteren, als het morgen zaterdag is?" (Bransford & Stein, 1984). Bij het oplossen van deze taken was het maken van een probleemdecompositie en een externe probleemrepresentatie noodzakelijk.

Het gebruik van externe representaties werd ook nagegaan in de dozen-toets, die bestond uit een aantal items, zoals "In een grote doos zitten twee middelgrote dozen. In de eerste middelgrote doos zitten drie kleine doosjes. In de tweede middelgrote doos zitten vier kleine doosjes. Hoeveel dozen zijn er in totaal?" (Bransford & Stein, 1984).

6 *Resultaten*

6.1 *Resultaten van de criteriumtoetsen*

Tabel 1 vermeldt voor elke criteriumtoets het aantal leerlingen in E1 en E2 dat het vooropgestelde beheersingscriterium (2/3 van de maximumscore) bereikte op het einde van het schooljaar.

Uit de resultaten op de toets over LOGO-kennis bleek dat op het einde van de leergang de meeste leerlingen uit beide experimentele klassen (21 en 17 in E1 resp. E2) een voldoende niveau van kennis van de LOGO-primitieven en concepten bereikt hadden (transfervoorwaarde 1).

De bevindingen op de toetsen betreffende de programmeerstrategie (transfervoorwaarde 2) waren eveneens erg positief. De resulta-

Tabel 1 *Aantal leerlingen in E1 en E2 dat het beheersingscriterium bereikte op de toetsen over de drie transfervoorwaarden*

Transfervoorwaarden	Criteriumtoetsen	Resultaten	
		E1 (n = 24)	E2 (n = 24)
1. LOGO-kennis	LOGO-kennis	21	17
2. Denkvaardigheden bij LOGO	Plannen 1	19	23
	Plannen 2	20	19
	Uitvoeren/controleren	20	22
3. Denkvaardigheden bij vraagstukken	Plannen 1	-	22
	Plannen 2	-	10
	Uitvoeren/controleren	-	22

ten van de eerste toets over de planfase toonden aan dat in E1 en E2 resp. 19 en 23 leerlingen voldoende in staat waren een reeks gegeven boomdiagrammen op hun juistheid te beoordelen. Uit de tweede toets over de planfase bleek dat in E1 en E2 resp. 20 en 19 leerlingen erin slaagden zelfstandig een adequaat boomdiagram te construeren voor een gegeven tekening. Bij de toets over de uitvoerings- en controlefase stelden we vast dat de aangeleerde werkwijze door de meeste leerlingen (20 en 22) nauwgezet gevolgd werd: bij de opbouw van hun programma gingen ze neerwaarts tewerk; de ingevoerde procedures werden regelmatig getest door het uitvoeren van de moederprocedure, en wanneer een verkeerd schermeffect optrad, verliep het opzoeken en verbeteren van fouten op een systematische manier. We kunnen dus stellen dat de beoogde denkvaardigheden, nl. plannen, debuggen, probleemdecompositie, en gebruik van een externe representatie, binnen de LOGO-context voldoende beheerst werden. Samenvattend kunnen we besluiten dat op het einde van de LOGO-cursus in beide experimentele klassen de eerste twee transfervoorwaarden vervuld waren.

Uit Tabel 1 blijkt verder dat de resultaten van de criteriumtoetsen in verband met de derde transfervoorwaarde, m.n. het kunnen toepassen van de geleerde vaardigheden bij het oplossen van andersoortige problemen, minder overtuigend waren. Immers, juist op de meest cruciale toets over deze voorwaarde, nl. deze waarbij de leerlingen zelfstandig de geleerde werkwijze voor het plannen moesten toepassen bij het oplossen van een reeks complexe vraagstukken (tweede toets), behaalden slechts 10 van de 24 leerlingen uit de tweede experimentele klas het criterium. Er was evenwel geen tijd meer beschikbaar voor remediëring, omdat het schooljaar op zijn einde liep. Omwille van deze laatste bevinding, was het meteen uitgesloten om op basis van de gegevens van dit onderzoek een beslissende toetsing te doen van het tweede gedeelte van de onderzoekshypothese, namelijk dat bij de verwezenlijking van de derde voorwaarde het transfereffect zal verhogen.

6.2 *Resultaten van de transfertoetsen*

Om de transfer van elk van de vier denkvaardigheden na te gaan werd een multivariate variantie-analyse (MANOVA) toegepast op de gegevens van de transfertoetsen. Afzonderlijke analyses werden uitgevoerd voor de vergelijking van E1 versus C enerzijds, en E1 versus E2 anderzijds. Volgens onze oorspronkelijke hypothese, zou E1 beter presenteren dan C, en zou E2 bovendien E1 overtreffen.

De multivariate analyse toonde een significant verschil tussen E1 en C op de maten voor debuggen, probleemdecompositie en externe representatie in het voordeel van E1 (zie Tabel 2). Hieruit kunnen we concluderen dat de LOGO-cursus een transfereffect teweeg bracht voor drie van de vier getrainde vaardigheden bij deze leerlingen bij wie de eerste twee transfervoorwaarden voldaan waren.

Voor de planvaardigheid werd geen significant verschil gevonden: beide groepen presteerden even goed op de toetsen die op deze

Tabel 2 *Vergelijking van E1 vs. C op de transfertoetsen (MANOVA)*

Denkvaardigheden	Transfertoetsen	Resultaten
Plannen	Doolhof, Foutendetectie	.249
Probleemdecompositie	Blokken, Weekdagen	*<.001*
Externe representatie	Weekdagen, Dozen	*<.001*
Debuggen	Foutendetectie	*.007*

vaardigheid beroep deden. Dit is enigszins verwonderlijk, daar de plancomponent de kern van de aangeleerde strategie uitmaakte. De meest voor de hand liggende verklaring hiervoor is naar onze mening dat de toets voor het meten van de planvaardigheid slechts een gebrekkige validiteit bezat.

Zoals verwacht kon worden op basis van het niet vervuld zijn van de derde transfervoorwaarde, maakte E2 niet meer vooruitgang dan E1.

7 *Samenvatting en discussie*

In onderhavige bijdrage werden de opzet en de resultaten beschreven van een onderwijsexperiment waarin we een LOGO-leeromgeving implementeerden, gericht op het verwerven en de transfer van algemeen bruikbare denkvaardigheden. Het onderzoek werd uitgevoerd in drie zesdeklassen, volgens een voortoetsnatoets opzet met controlegroep.

Op basis van de recente literatuur formuleerden we drie voorwaarden die van belang zijn voor het bereiken van transfer van denkvaardigheden. De eerste twee voorwaarden – het verwerven van een voldoende niveau van domeinspecifieke kennis enerzijds, en de beheersing van de beoogde denkvaardigheden binnen de programmeercontext anderzijds – trachtten we in beide experimentele klassen te realiseren in een LOGO-cursus. In de tweede experimentele klas werd daarenboven expliciet transfer nagestreefd (derde voorwaarde) in een bijkomende vraagstukkencursus.

Alvorens het transfereffect na te gaan, onderzochten we of de nagestreefde *transfervoorwaarden* vervuld waren. Uit de resultaten van de criteriumtoetsen over LOGO bleek dat op het einde van de leergang de eerste twee voorwaarden in beide experimentele klassen voldoende gerealiseerd waren. De meeste leerlingen hadden de aangeleerde primitieven en concepten van de LOGO-taal goed verworven en waren in staat de geleerde denkvaardigheden toe te passen binnen de LOGO-context. Dat deze gunstige uitkomsten werden bereikt, kan onzes inziens toegeschreven worden aan de zorgvuldig ontwikkelde onderwijsleeromgeving, waarin tegemoet gekomen werd aan de tekortkomingen van de vroegere studies. Twee belangrijke kenmerken van deze leeromgevingen waren: 1) de relatief lange duur van de LOGO-instructie en 2) het evenwicht tussen zelfontdekkend leren en systematische instructie (bijv. het expliciet aanleren van een programmeerstrategie). Toch willen we waarschuwen voor een al te groot optimisme omtrent de veralgemeenbaarheid van deze resultaten. Men mag namelijk niet uit het oog verliezen dat de gehanteerde leeromgeving allesbehalve representatief is voor de feitelijke situatie in de meeste lagere scholen anno 1989: er werd gedurende 60 uur LOGO-instructie gegeven door drie ervaren begeleiders in samenwerking met de leerkracht en daarbij waren negen computers beschikbaar. En zelfs dan nog bleven bij verscheidene leerlingen lacunes en hardnekkige misvattingen bestaan, die slechts via geïndividualiseerde remediëring verholpen konden worden.

We slaagden er niet in de derde transfervoorwaarde te realiseren. Zelfs na een periode van expliciete instructie in het aanwenden van de in de LOGO-context geleerde probleemoplossingsvaardigheden bij het oplossen van complexe rekenvraagstukken, bleken de meeste leerlingen in E2 onvoldoende in staat deze vaardigheden zelfstandig toe te passen bij de vraagstukken uit de betreffende criteriumtoets. Deze teleurstellende resultaten zijn waarschijnlijk toe te schrijven aan de stevig verankerde, maar eerder ongunstige aanpakwijzen en denkbeelden in verband met wiskundige problemen in het algemeen en met vraagstukken in het bijzonder, die bij deze leerlingen vooraf reeds bestonden ten gevolge van jarenlang vraagstukkenonderwijs (Sowder, 1988). Allicht volstonden de duur en de

aard van de instructie niet om wezenlijke veranderingen op strategisch vlak te bewerkstelligen.

Wat het eigenlijke *transfereffect* betreft, werden bemoedigende resultaten verkregen. Bij de leerlingen die de LOGO-cursus volgden (E1), werd voor drie van de vier gemeten vaardigheden, nl. debugging, decompositie, en externe representatie, transfer bereikt. Inmiddels zijn in andere recente studies gelijkaardige positieve resultaten geboekt (zie o.m. Mayer, 1988; voor een overzicht, zie De Corte & Verschaffel, 1987, ter perse). Een gemeenschappelijke factor in deze succesrijke studies is de expliciete instructie van de beoogde denkvaardigheden in de LOGO-context, waarbij deze vaardigheden bewust en intentioneel geabstraheerd worden; Salomon en Perkins (1987) noemen dit 'mindful abstraction'.

Het tweede deel van onze verwachting, m.n. dat de expliciete instructie voor transfer het effect zou verhogen, werd niet ingelost: de leerlingen die bijkomend een vraagstukkencursus gevolgd hadden (E2), behaalden op geen van de vier maten hogere resultaten dan deze uit E1. Toch zou het voorbarig zijn om hieruit af te leiden dat het expliciet nastreven van transfer naar verschillende domeinen geen gunstige invloed heeft op de algemene denkontwikkeling van leerlingen. Immers, in onderhavige studie kan het ontbreken van het verwachte transfereffect verklaard worden door het niet bereiken van de doelstelling van de vraagstukkencursus, m.n. het leren toepassen van de denkvaardigheden verworven in de LOGO-context bij het oplossen van complexe rekenvraagstukken (= transfervoorwaarde 3). Daardoor kwam het toetsen van de hypothese in het kader van deze studie als het ware al bij voorbaat op de helling te staan. Mede op grond van recente literatuur (o.m. Gick & Holyoak, 1987; Perkins & Salomon, 1989), blijven we evenwel de hypothese aanhouden dat, indien men erin slaagt de derde voorwaarde te realiseren, er wel een verhoging van het transfereffect kan verwacht worden. Overigens werd in enkele meer recente studies (zie o.m. Littlefield, Delclos, Lever, Clayton, Bransford & Franks, 1988) wel steun voor deze hypothese verkregen.

Het uitgangspunt van onderhavig onderzoek was de vraag "Heeft het leren programmeren een invloed op de denkvaardigheden van basisschoolleerlingen?". Op basis van de verkregen onderzoeksresultaten en de opgedane ervaringen menen we hierop een gematigd positief antwoord te kunnen formuleren. *Ja,* de bevindingen suggereren dat leren programmeren in LOGO kan leiden tot het verwerven en de transfer van algemene denkvaardigheden. *Maar,* de positieve effecten vloeien allicht niet louter voort uit de unieke eigenschappen van LOGO als dusdanig, maar eerder uit de kwaliteiten van de onderwijsleeromgeving waarin de programmeertaal aangeleerd wordt. Met andere woorden, LOGO *op zich* biedt geen garantie voor het tot stand komen van denkvaardigheden, maar kan hiervoor een nuttig hulpmiddel zijn als het ingebed is in een krachtige leeromgeving, d.w.z. een situatie die aanleiding geeft tot het uitlokken bij de leerlingen van de vereiste leerprocessen om de beoogde onderwijsdoelen te bereiken.

Voortgezet onderzoek zou gericht moeten zijn op het meer precies identificeren en analyseren van de cruciale kenmerken van zulke leeromgevingen. Een sleutelvraag in dit verband is of, en in hoeverre de noodzakelijke aanvullende denkinstructie het best in de programmeeromgeving zélf wordt ingebouwd, en welk gedeelte beter in handen van de leerkracht kan blijven. Met andere woorden, moet er gewerkt worden in de richting van intelligente systemen die op basis van verfijnde diagnoses van de vaardigheden en aanpakstrategieën van de leerlingen aangepaste begeleiding kunnen voorzien? Of moeten deze activiteiten eerder uitgevoerd worden door de leerkracht en blijft de rol van de computer het best beperkt tot het bieden van een situatie die de leerlingen uitnodigt tot het maximaal aanwenden van hun eigen kennis en denkvermogens (Scardamalia, Bereiter, McLean, Swallow, & Woodruff, 1989)? Een voorbeeld van een systeem waarin de oriëntering naar algemene denkvaardigheden sterker door de computer zelf ondersteund wordt, is het Michaelsysteem, dat ontwikkeld werd door de Afdeling Informatica van het Departement Computerwetenschappen van de K.U. Leuven, in samenwerking met onze onderzoeksgroep. Dit systeem stimuleert uitdrukkelijk een stapsgewijze opbouw en verfijning van programma's. Tevens biedt het verschillende mogelijkheden om de gebruiker te bevrijden van de syntactische aspecten van het programme-

ren, waardoor de aandacht kan toegespitst worden op de meer essentiële probleemoplossings- en ontwerpactiviteiten (Tollenaere, 1988).

Tot slot lijkt het ons aangewezen het verder onderzoeks- en ontwikkelingswerk niet te beperken tot programmeer-leeromgevingen. Aangezien de vastgestelde cognitieve effecten allicht geen rechtstreeks resultaat zijn van het programmeren op zich, maar eerder van de gehele leeromgeving waarin dit ingebed wordt, kunnen we aan het leren programmeren geen monopoliepositie inzake denktraining toekennen. Bijgevolg mogen we veronderstellen dat, gegeven de investering van eenzelfde hoeveelheid tijd, inspanning en expertise, even efficiënte leeromgevingen kunnen gecreëerd worden in meer 'traditionele' inhoudsdomeinen; dit zou daarenboven de integratie van de denktraining in het reguliere curriculum ten goede komen. Hiermee willen we pleiten voor de aanpak van de denkscholing vanuit verschillende inhoudsdomeinen. Overigens vereist de totstandkoming van *algemeen bruikbare* vaardigheden dat de kinderen deze leren toepassen in verschillende vakgebieden en probleemsituaties. Dit kan enkel verwezenlijkt worden indien er vakoverschrijdend gepland wordt hoe men het transfergericht leren vorm zal geven.

Noten

1. Bij de tussenstukken geldt de volgende afspraak. De naam van een tussenstuk bestaat uit twee delen: ts (van tussenstuk), gevolgd door de eerste letter van het vorige deel en de eerste letter van het volgende deel (bijv.: *t*ussens*t*uk *v*enster *d*ak: ts-vd).
2. Om overzichtelijke programma's te bekomen raden we de kinderen aan om een nieuwe lijn te beginnen voor iedere procedure in een superprocedure. In deze tekst staan echter wegens plaatsgebrek alle procedures op één lijn.

Literatuur

Bransford, J. D. & B. S. Stein, *The ideal problem solver.* New York: Freeman, 1984.

Collins, A., J. S. Brown & S. E. Newman, Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. In: L. B. Resnick (Ed.), *Cognition and instruction: Issues and agendas.* Hillsdale, NJ: Erlbaum, in press.

Corte, E. De, Ontwerpen van krachtige onderwijsleeromgevingen. In: M. J. Ippel & J. J. Elshout (Eds.), *Training van hogere-orde denkprocessen.* Lisse: Swets & Zeitlinger, 1990.

Corte, E. De & L. Verschaffel, Cognitieve effecten van leren programmeren. *Nederlands Tijdschrift voor de Psychologie,* 1987, *42,* 364-372.

Corte, E. De & L. Verschaffel, Leuvense perspectieven: Informatica en leren probleemoplossend denken in het onderwijs. *Onze Alma Mater,* ter perse.

Corte, E. De, L. Verschaffel, E. Hoedemaekers, R. Indemans & H. Schrooten, *Computers en Leren Denken. Activiteitenverslag over het werkjaar 1987.* (Project "Computers en Leren Denken", Rapport nr. 6). Leuven: Onderzoekscentrum voor Onderwijsleerprocessen, K.U. Leuven, 1988.

Corte, E. De, L. Verschaffel, H. Schrooten, R. Indemans & E. Hoedemaekers, *LOGO als springplank voor het verwerven van denkvaardigheden bij zesdeklassers* (Leuvense Pedagogische Bijdragen). Leuven: Acco, 1990.

Gick, M. L. & K. J. Holyoak, The cognitive basis of knowledge transfer. In: S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning. Contemporary research and applications.* (Educational Technology Series). San Diego: Academic Press, 1987.

Littlefield, J., V. R. Delclos, S. Lever, K. N. Clayton, J. D. Bransford & J. J. Franks, Learning LOGO: Method of teaching, transfer of general skills, and attitude toward school and computers. In: R. E. Mayer (Ed.), *Teaching and learning computer programming. Multiple research perspectives* (pp. 111-135). Hillsdale, NJ: Erlbaum, 1988.

Mayer, R. E. (Ed.), *Teaching and learning computer programming. Multiple research perspectives.* Hillsdale, NJ: Erlbaum, 1988.

McCoy Carver, S., *Transfer of LOGO debugging skill: Analysis, instruction and assessment.* (Doctoral dissertation). Pittsburgh, PA: Department of Psychology, Carnegie-Mellon University, 1986.

Papert, S., *Mindstorms. Children, computers, and powerful ideas.* New York: Basic Books, 1980.

Papert, S., *Computers en kinderen.* Amsterdam: Bert Bakker, 1984.

Perkins, D.N. & G. Salomon, Are cognitive skills context-bound? *Educational Researcher,* 1989, *18,* no. 1, 16-25.

Salomon, G. & D.N. Perkins, Transfer of cognitive skills for programming: When and how? *Journal of Educational Computing Research,* 1987, *3,* 149-169.

Scardamalia, M., C. Bereiter, R.S. McLean, J. Swallow & E. Woodruff, Computer-supported intentional learning environments, *Journal of Educational Computing Research,* 1989, *5,* 51-68.

Sowder, L., Children's solutions of story problems. *Journal of Mathematical Behavior,* 1988, *7,* 227-238.

Tollenaere, T., *Een didactische programmeeromgeving* (niet-gepubliceerde licentiaatsverhandeling). Leuven: K.U. Leuven, Departement Computerwetenschappen, 1988.

Verschaffel, L., E. De Corte, E. Hoedemaekers, H. Schrooten & R. Indemans, *Beheersing van de LOGO-kennis bij zesdeklassers na een jaar LOGO-instructie.* (Project "Computers en Leren Denken", Rapport nr. 5). Leuven: Onderzoekscentrum voor Onderwijsleerprocessen, K.U. Leuven, 1987.

Curricula vitae

E. De Corte is gewoon hoogleraar aan de Faculteit der Psychologie en Pedagogische Wetenschappen, K.U. Leuven.

L. Verschaffel is bevoegdverklaard navorser bij het Nationaal Fonds voor Wetenschappelijk Onderzoek en buitengewoon docent aan de Faculteit der Psychologie en Pedagogische Wetenschappen, K.U. Leuven.

H. Schrooten is wetenschappelijk medewerker in het onderzoeksproject "Computers en Leren Denken" van het Onderzoekscentrum voor Onderwijsleerprocessen, Faculteit der Psychologie en Pedagogische Wetenschappen, K.U. Leuven.

Adres: Onderzoekscentrum voor Onderwijsleerprocessen, Departement der Psychologie en Pedagogische Wetenschappen, K.U. Leuven, Vesaliusstraat 2, B-3000 Leuven, België.

Manuscript aanvaard 12-3-90

Summary

Corte, E. De, L. Verschaffel & H. Schrooten. 'Cognitive effects of learning to program in LOGO: A one-year study with sixth graders.' *Pedagogische Studiën,* 1990, *67,* 293-305.

This study aimed at the development of a powerful LOGO learning environment for achieving the following conditions for the acquisition and transfer of thinking skills: 1) sufficient knowledge of the domain specific LOGO primitives and concepts; 2) mastery of the thinking skills within the programming context; and 3) ability to apply these skills in at least one other content domain. A systematic experiment was carried out in three sixth grade classes according to the pretest-posttest design with control group. A LOGO course, involving the training of a systematic programming strategy, was taught in two experimental classes. In one of these classes the pupils also learned how to apply the acquired skills in another content domain. The control group was a non-treatment group. At the end of the school year condition fulfilment was tested. The findings showed that the firtst two transfer conditions were fulfilled in both experimental groups; the results with respect to the third condition were less positive. Furthermore, a series of transfer tests were administered. Data analysis revealed that transfer was obtained in both experimental groups. This implies that fulfilment of the first two conditions mentioned above is sufficient for realizing transfer.

Rekenfeiten in het cijferend optellen en aftrekken tot 100*

M. W. J. BALTUSSEN en E. C. D. M. VAN LIESHOUT**
Instituut voor Orthopedagogiek
Katholieke Universiteit Nijmegen

Samenvatting

De vraag naar de functionele rol van het automatiseren van de rekenfeiten bij het cijferend optellen en aftrekken staat in dit onderzoek centraal. Bij 58 lom- en ml-kinderen is nagegaan of beheersing van de rekenfeiten direct van invloed is op de juistheid en het tempo van oplossen. Eveneens is onderzocht of deze beheersing indirect via het gebruik van geavanceerde rekenstrategieën in de juistheid van het antwoord tot uitdrukking komt. Er kon geen steun gevonden worden voor de hypothese dat kinderen met meer rekenfeitenkennis bij het oplossen van cijferopgaven minder vertraagd zouden worden door tientaloverschrijding dan kinderen zonder rekenfeitenkennis. Het ontbreken van rekenfeitenkennis blijkt wel vertragend te werken bij opgaven waarbij een twee-cijferig getal opgeteld of afgetrokken moet worden. Daarnaast bleken kinderen zonder kennis van de rekenfeiten vaker een telstrategie toe te passen en meer telfouten te maken dan kinderen die over meer geautomatiseerde rekenfeiten beschikken. Er lijken aanwijzingen te zijn voor een functionele rol van de rekenfeiten bij het cijferend rekenen.

* Deze bijdrage vormt een bewerkte versie van het paper: "Number facts in het cijferend optellen en aftrekken tot 100" gepresenteerd tijdens de Onderwijsresearchdagen, 24-25 mei 1989 te Leiden.

** Wij willen de teams van de Zonnegaardschool en de Wilhelmina Bladergroenschool uit Nijmegen en De Schans uit Venlo bedanken voor hun medewerking aan dit onderzoek. Tevens dank aan J. van Erp en J. van Veldhoven voor hun assistentie bij de onderzoekswerkzaamheden.

1 *Inleiding*

Kinderen met rekenproblemen (afkomstig uit het gewoon lager onderwijs) rekenen op een niveau dat vergelijkbaar is met normaal lerende maar een jaar jongere kinderen (Russel & Ginsburg, 1984). Een uitzondering vormt de beheersing van de rekenfeiten, die meer dan een jaar achterblijft. De rekenfeiten zijn onvoldoende geautomatiseerd. Dit zou een mogelijke oorzaak van rekenproblemen kunnen zijn. Onder rekenfeiten verstaan wij in dit onderzoek alle optelopgaven met enkelvoudige getallen zoals $7+8$ en aftrekopgaven met enkelvoudig aftrektal en enkelvoudige uitkomst zoals $12-5$. Wij zullen ons in deze bijdrage beperken tot die rekenfeiten waarbij sprake is van tientaloverschrijding. In onderhavig onderzoek wordt nagegaan of er aanwijzingen zijn dat de beheersing van rekenfeiten van belang is voor prestaties op het gebied van het cijferend rekenen. Wij richten ons hierbij op kinderen uit het speciaal onderwijs.

In een leerproces treden na verloop van tijd verkorting en versnelling van procedures op: handelingen worden geautomatiseerd. Automatisering van handelingen betekent echter niet alleen dat deze handelingen verkort en sneller uitgevoerd gaan worden, ook de graad van bewustzijn van de handelingen verandert. De procedure hoeft niet meer bewust uitgevoerd te worden maar voltrekt zich automatisch (Laberge & Samuels, 1974; Eysenck, 1984). Deze ideeën zijn met name ontwikkeld op het terrein van het lezen. Indien een geoefende lezer geconfronteerd wordt met een woord dan zal hij dit spontaan lezen; hij kan dit zelfs niet voorkómen. Bij gebrek aan vergelijkbare theorieën voor het rekenen, wordt verondersteld dat de theorieën omtrent de geautomatiseerde woordherkenning enigszins van toepassing zijn op de rekenfeiten (Resnick & Ford, 1984). Of ook hier geldt dat het zien van de opgave $6+5$, spontaan het antwoord '11' doet opkomen, is uit de huidige stand van onderzoek niet af te leiden. Wel lijkt bij een geautomatiseerd zijn van de rekenfeiten, evenals bij de geautomatiseerde woordherkenning de uitvoer van handelingen onbewust te verlopen zodat geen beroep meer gedaan

hoeft te worden op het werkgeheugen. De beperkte geheugencapaciteit kan daardoor voor andere – niet geautomatiseerde – zaken ingezet worden (Perfetti & Lesgold, 1979; Anderson, 1982).

Kinderen met rekenproblemen blijven rekenfeiten echter veelal langzaam en niet foutloos oplossen. Aanvankelijk bereiken de meeste kinderen de oplossing op een rekenfeit via uitgebreide procedures. Dit kan via het geheel uittellen van het antwoord (count-all) of via het geavanceerde doortellen (count-on) of bijtellen van het kleinste getal (count-by-min) verlopen (Groen & Parkman, 1972; Fuson, 1982). Uiteindelijk zullen de procedures zodanig verkort en versneld zijn dat het antwoord ineens uit het geheugen opgehaald kan worden. Het doorlopen van telprocedures blijft achterwege. De antwoorden op de rekenfeiten liggen direct toegankelijk opgeslagen in het lange termijn geheugen (Ashcraft & Battaglia, 1978; Ashcraft, 1982; Hamann & Ashcraft, 1985). De uitgebreide telprocedures blijven aanwezig zodat bij het verloren gaan van de direct uit het geheugen op te halen informatie hierop teruggevallen kan worden (Payne, 1988).

Naast het ineens ophalen van het antwoord uit het geheugen kunnen rekenfeiten eveneens opgelost worden via het gebruik van afgeleide rekenfeiten (Wolters, 1984/1985; Baroody, 1985; Steinberg, 1985). Zogenaamde handige rekenregels bieden de mogelijkheid om een onbekend rekenfeit om te vormen tot een vertrouwd probleem. Bijvoorbeeld via de dubbel + 1-regel (Steinberg, 1985) kan de som 6+7 herleid worden tot de bekende som (6+6)+1. De handige rekenregels zijn specifiek afgestemd op één somtype. Het gevaar voor overgeneralisatie is met name bij zwakkere rekenaars groot (Wolters, 1984/1985). De kans op fouten in de rekenfeiten is bij deze groep kinderen waarschijnlijk kleiner indien het antwoord direct uit het geheugen opgehaald kan worden.

Het geautomatiseerd beheersen van de rekenfeiten heeft daarnaast, zoals eerder vermeld, als verdienste dat het werkgeheugen minder belast wordt in vergelijking met het uittellen of afleiden van het antwoord. Er blijft ruimte over voor het aanwenden van niet-geautomatiseerde procedures zoals bijvoorbeeld het inwisselen (Anderson, 1982; Resnick & Ford, 1984). Dit leidt mogelijk tot een vermindering van het foutenaantal.

Zoals eerder gesteld blijven kinderen met rekenproblemen achter wat betreft snelheid en nauwkeurigheid bij het oplossen van rekenfeiten. Het grote aantal van buiten te leren rekenfeiten (100) zou mogelijk voor deze kinderen een te zware geheugenbelasting vormen. In recent onderzoek (Baltussen & Van Lieshout, 1989) werd daarom het belang van het geautomatiseerd zijn van de splitsingen voor de cijferprestaties nagegaan. Splitsen omvat het uiteen kunnen leggen van een getal (tot en met 10) in zijn samenstellende delen. Bijvoorbeeld 5 kan gesplitst worden in 2 en 3 of 1 en 4. Een splitsing wijkt qua vorm af van een rekenfeit. Heeft een rekenfeit de vertrouwde vorm van een optel- of aftreksom, de splitsing wordt schriftelijk veelal weergegeven in een T-vorm: $\frac{8}{3\ 1}$. In totaal zijn er 66 splitsingen.

Verwacht werd dat het geautomatiseerd zijn van de splitsingen, het oplossen van sommen met tientaloverschrijding zou vergemakkelijken doordat splitsen aanleiding geeft tot het gebruik van een splitsstrategie (aanvullen en leegmaken tot 10); bijvoorbeeld bij: 7+8 dat splitsend uitgerekend wordt als (7+3) + 5. Het gebruik van splitsingen bij het oplossen van opgaven met tientaloverschrijding betekent een reductie in het aantal van buiten te leren rekenfeiten. In het voorbeeld hoeft het rekenfeit 7 + 8 = 15 niet meer gekend te worden. Wel moet het inzetten van de splitsstrategie aangeleerd worden. Dit algoritme is vergelijkbaar voor alle cijferopgaven en zou derhalve minder geheugenruimte behoeven.

In het genoemde onderzoek kon echter de interactie tussen de splitsvaardigheid van het kind en de aanwezigheid van tientaloverschrijding noch in het aantal juiste oplossingen noch in de oplossingstijd aangetoond worden. Dit leidde tot de vraag of de beschikbaarheid van de rekenfeiten met tientaloverschrijding wellicht toch een belangrijke rol vervult tijdens het cijferen van kinderen met rekenproblemen. Het zelfstandig toepassen van de splitsstrategie vormt mogelijk een groter struikelblok voor deze kinderen dan het van buiten leren van een grotere hoeveelheid rekenfeiten. Wellicht leidt het niet beschikbaar hebben van de rekenfeiten met tientaloverschrijding eerder tot tellen dan tot spontaan toepassen van een splitsstrategie zodat

het tempo vertraagt en het aantal fouten toeneemt. De gegevens uit het genoemde onderzoek werden naar aanleiding van deze vraag heranalyseerd. Het achterliggende idee bij deze vraagstelling wordt gevormd door de overweging dat met name kinderen met rekenproblemen moeite blijven houden met het automatiseren van de rekenfeiten. Is het noodzakelijk deze kinderen te blijven trainen op dit terrein of kan ondanks het niet beheersen van de rekenfeiten toch verder gegaan worden met het voortgezet rekenen terwijl een redelijke kans op succes behouden blijft?

Het toepassen van rekenfeiten met tientaloverschrijding is vooral nuttig bij opgaven met tientaloverschrijding. De eenheden kunnen zonder tussenstappen opgeteld of afgetrokken worden. Kinderen, die deze rekenfeiten geautomatiseerd tot hun beschikking hebben, zullen bij opgaven met tientaloverschrijding in het voordeel zijn ten opzichte van kinderen die niet beschikken over geautomatiseerde rekenfeiten. De laatst genoemde kinderen zullen extra tijd nodig hebben voor het in het geheugen zoeken, het uitrekenen of het uittellen van het antwoord of, nadat ze in een impasse zijn geraakt, voor het overgaan op een andere oplossingsstrategie. Verwacht wordt, dat kinderen, die de rekenfeiten geautomatiseerd hebben, door de aanwezigheid van tientaloverschrijding bij het oplossen van cijferopgaven minder vertraagd worden dan kinderen zonder geautomatiseerde kennis van de rekenfeiten.

Een vrijwel identieke voorspelling kan, hoewel minder gemakkelijk, ook voor de juistheid van oplossen van cijferopgaven gedaan worden. Een goed antwoord is weliswaar ook via tellen, splitsen of handig rekenen te verkrijgen, maar de kans op fouten lijkt groter als gevolg van onder andere de zwaardere geheugenbelasting. Bij tellen is daarnaast het dubbelsporig tellen nog een extra foutenbron. Onder dubbelsporig tellen wordt verstaan het tegelijkertijd bijhouden van hoeveel al bij- of afgeteld is en van de tussenuitkomst op dat moment. Bijv. $8+4=9\,(+1)$, $10\,(+2)$, $11\,(+3)$, $12\,(+4)$. Het gebruik van rekenfeiten kan in eerste instantie, door verkeerd onthouden, leiden tot een toename van het aantal fouten. Zijn de rekenfeiten echter volledig geautomatiseerd dan zal het foutenaantal waarschijnlijk dalen.

Een interactie-effect tussen tientaloverschrijding en het beheersen van rekenfeiten voor wat betreft het aantal juist opgeloste opgaven zou derhalve een extra bevestiging voor de rol van de rekenfeiten zijn. Kinderen, die beschikken over geautomatiseerde kennis van de rekenfeiten zullen relatief meer opgaven met tientaloverschrijding goed maken dan kinderen die de rekenfeiten niet geautomatiseerd tot hun beschikking hebben, terwijl ze minder van elkaar zullen verschillen bij opgaven zonder tientaloverschrijding.

Observatie van oplossingsgedrag kan een andere aanwijzing geven voor de functionele rol, die het beheersen van de rekenfeiten bij het cijferend rekenen speelt. Kennis van de rekenfeiten is functioneel als deze bij het oplossen van cijferopgaven ingezet wordt. Hoewel veel kinderen vaak uit veiligheidsoverwegingen blijven tellen, sluipen in deze oplossingsroute fouten, die door het automatiseren van de rekenfeiten voorkomen kunnen worden. Kinderen, die de rekenfeiten geautomatiseerd hebben, zullen naar verwachting geavanceerdere oplossingsstrategieën (met name een rekenfeitenstrategie) gebruiken dan kinderen die de rekenfeiten niet geautomatiseerd hebben. Bevestiging van deze hypothese leidt tot de vraag of het gebruik van een geavanceerde oplossingsstrategie betere prestaties tot gevolg heeft. Kinderen, die tijdens het uitrekenen van een cijfersom een geavanceerde oplossingsstrategie gebruiken, zullen meer goede antwoorden geven dan kinderen die zich bedienen van minder geavanceerde strategieën. Een positief antwoord zou een verdere aanwijzing zijn voor de rol van de rekenfeiten bij het cijferend rekenen.

Het analyseren van foutief opgeloste opgaven biedt een andere mogelijkheid om informatie te verkrijgen over het inzetten van rekenfeiten bij het oplossen van cijferopgaven. Kinderen, die de rekenfeiten beheersen, zullen weinig telfouten maken. Hun fouten zullen zich concentreren in fouten tegen nog niet geautomatiseerde procedures zoals inwisselen en lenen. Kinderen, die de rekenfeiten niet geautomatiseerd hebben daarentegen, zullen naar verwachting zowel telfouten als fouten met inwisselen en lenen maken. Foutenanalyse kan meer inzicht geven in het daadwerkelijk gebruik van rekenfeiten.

2 *Methode*

2.1 *Procedure*

Het onderzoek viel uiteen in twee delen. Het eerste deel bestond uit een drietal individuele schriftelijke en mondelinge toetsen, die eenduidig te scoren waren. Het tweede deel bevatte één toets die aan de hand van hardopdenkonderzoek de door de leerling gehanteerde oplossingsstrategieën trachtte te achterhalen. Uit betrouwbaarheidsoverwegingen was bij het hardopdenk-onderzoek een tweede observator aanwezig. Voor het tweede deel van het onderzoek werd, vanwege de doelstellingen van het oorspronkelijke onderzoek waarin het splitsen centraal stond (Baltussen & Van Lieshout, 1989), een selectie uit de totale onderzoekspopulatie gemaakt op basis van splitsvaardigheid (de gemiddelde oplossingstijd per goede splitsopgave). Splitsvaardigheid kende twee categorieën a) hoog: de gemiddelde oplossingstijd $\leq$ 2.00 seconden en b) laag: gemiddelde oplossingstijd $>$ 2.00 seconden. De korte oplossingstijd doet een geautomatiseerd zijn van de splitsingen vermoeden. Verondersteld werd dat de kinderen zonder bewust te hoeven nadenken, de opgaven oplossen. Dit kan echter niet in dit onderzoek nagegaan worden. Bij de hier beschreven heranalyse is een selectie gemaakt uit de kinderen op basis van de beheersing van de rekenfeiten. Daar van de totale onderzoeksgroep slechts twee kinderen de rekenfeiten in gemiddeld minder dan 2 seconden goed oplosten, werd gekozen voor het aanpassen van het tijdscriterium. De literatuur biedt geen duidelijke definitie van automatisering van de rekenfeiten. De vereiste oplossingstijd varieert tussen de 4 en 5 seconden (Usnick & Engelhardt, 1988). Dit leek ons echter te hoog. In 4 seconden kan een opgave tellend opgelost worden. Om toch twee groepen te formeren die duidelijk verschillen in de mate van geautomatiseerd zijn van de rekenfeiten is voor de snelle groep als criterium aangehouden: tenminste vijf van de tien rekenfeiten in $\leq$ 2 seconden goed opgelost. Voor de langzame groep was de maatstaf: geen enkel rekenfeit in $\leq$ 2 seconden goed beantwoord. Dit betekent dat de kinderen in de snelle groep niet alle rekenfeiten geautomatiseerd hebben. Aangenomen werd dat de graad van automatisering desondanks hoger was in deze groep dan in de groep met een gemiddeld langere antwoordtijd.

2.2 *Materiaal*

In het eerste gedeelte van het onderzoek werd de proefpersonen een splitstoets bestaande uit 15 splitsingen en een rekenfeitentoets bestaande uit tien opgaven betrekking hebbend op rekenfeiten met tientaloverschrijding (5 optelen 5 aftrekopgaven) voorgelegd. De antwoordtijd en juistheid werden gescoord. Daarnaast werd een cijfertoets afgenomen, die bestond uit 32 cijferopgaven waarvan de ene helft optel- en de andere helft aftrekopgaven bevatte (zie Bijlage 1). De optel- en aftrekopgaven waren verdeeld in acht opgaven zonder en acht opgaven met tientaloverschrijding. Uitgaand van eerdere onderzoeksresultaten (Andrea, 1988) werd verondersteld dat zeer zwakke rekenaars het tweede getal uit een opgave geheel tellend bij het eerste getal zouden optellen of aftellen. De grootte van de telafstand leek derhalve ook de moeilijkheidsgraad van de opgave te beïnvloeden. De groepen van acht opgaven met en zonder tientaloverschrijding werden daarom op hun beurt verdeeld in drie opgaven met een kleine telafstand (+ of − max. 9) en vijf opgaven met een grote telafstand (+ of − min. 21). Dit alles resulteerde in acht verschillende somtypen. De antwoorden op de cijfertoets werden gescoord op tijd en juistheid.

Het hardopdenk-onderzoek bestond uit een strategie-onderzoek waarin het kind hardop denkend de opgaven moest oplossen. Het strategie-onderzoek bevatte 5 optel- en 5 aftrekopgaven met tientaloverschrijding (zie Bijlage 2). De scoring van de strategieën verliep aan de hand van een categorieënsysteem op basis van de methode die het kind hanteerde voor het optellen en aftrekken van de eenheden. De mate van geavanceerdheid van een oplossingsstrategie werd hierbij afgeleid uit de efficiëntie qua geheugenbelasting en de juistheid van de strategie. Onderscheiden werden de volgende strategieën in oplopende graad van geavanceerdheid: verkeerde strategie (smaller from larger, verkeerd splitsen en alle cijfers optellen alsof het eenheden zijn), tellen, splitsen, handig rekenen en/of rekenfeit toepassen. De geavanceerdheid van het strategiegebruik werd gescoord in een 4-puntsschaal (zie Figuur 1).

De categorieën, die geen relevante informatie over het optellen of aftrekken van de eenheden leverden, werden buiten beschouwing gelaten. Van 2.1% van de opgaven was

score 0:	verkeerde strategie		
	– foutief splitsen:	36	
		+18	8 gesplitst in 4 en 4
		54	
	– smaller from larger:	43	
		–17	7–3=4
		34	
	– alles optellen:	36	
		+27	3+6+2+7=18
		18	
score 1:	telstrategie	36	
		+17	6+7=6,7,8...,13
		53	
score 2:	splitsstrategie	43	
		–17	eerst 3 eraf, inwisselen,
			daarna nog 4 eraf
		26	
score 3:	rekenfeitstrategie	43	
		–17	13–7=6
		26	
	handig rekenstrategie	36	
		+17	6+7=(6+6)+1=12+1=13
		53	

Figuur 1 *Scoringsschaal ten behoeve van het strategiegebruik*

het strategiegebruik niet categoriseerbaar.

Per persoon werd per opgave de strategie bepaald, gescoord op de schaal (van 0 t/m 3) en de totale score over de 10 opgaven berekend. Dit leidde uiteindelijk tot een strategievaardigheidsscore variërend van 0 t/m 30.

Binnen de categorieën tel-, rekenfeit- en handig rekenstrategie is geen onderscheid gemaakt tussen het juist of onjuist gebruik van deze strategieën. In eerste instantie omdat het oorspronkelijke onderzoek gericht was op de vergelijking van met name goede en slechte splitsers. Toen echter bleek dat ook kinderen uit het speciaal onderwijs een breed scala aan oplossingsstrategieën tot hun beschikking hadden, werd na bestudering van het aantal foutief toegepaste rekenfeit-, handig reken- en telstrategieën besloten dat een dergelijk onderscheid met het oog op het geringe voorkomen van foutief strategiegebruik onnodig was.

2.3 *Proefpersonen*

Voor het oorspronkelijke onderzoek werden 58 leerlingen van twee lom-scholen en een mlk-school geselecteerd. Er werden zoveel mogelijk kinderen geselecteerd, die nog niet zover in het cijferonderwijs gevorderd waren dat ze het oplossen van cijfersommen volledig onder de knie hadden. De uiteindelijke onder-

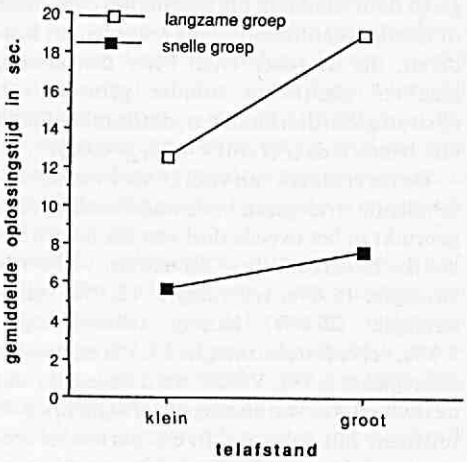
zoeksgroep voor het eerste deel van het onderzoek bestond uit 39 kinderen uit het lom-onderwijs (30 jongens, 9 meisjes; gemiddelde leeftijd: 10.7 jaar, $sd = 0.95$) en 19 kinderen uit het ml-onderwijs (9 jongens, 10 meisjes; gemiddelde leeftijd: 12.4 jaar, $sd = 1.01$). Voor het tweede deel van het onderzoek werden van elke school de vijf kinderen met de hoogste en de vijf kinderen met de laagste splitsvaardigheid gekozen.

Bij de heranalyse werd volgens het criterium genoemd in 2.1 uit de genoemde proefgroep een nieuwe populatie samengesteld voor de herhaalde metingen analyse. Deze populatie bestond uit 20 kinderen uit het lom-onderwijs (14 jongens, 6 meisjes; gemiddelde leeftijd: 10;8 jaar, $sd = 0.95$) en 9 kinderen uit het ml-onderwijs (3 jongens en 6 meisjes; gemiddelde leeftijd: 12;6 jaar, $sd = 1.13$). Deze procedure resulteerde in acht kinderen voor de snelle en tien kinderen voor de langzame groep in het tweede deel van het onderzoek.

3 *Resultaten*

Oplossingstempo

In een $2 \times 2 \times 2 \times 2$ variantie analytisch design met als afhankelijke variabele het oplossingstempo van de cijferopgaven, als tussen subjecten factor: beheersing van de rekenfeiten en met herhaalde metingen over de factoren: operatie, tientaloverschrijding en afstand werd geen significant interactie-effect gevonden tussen beheersing van de rekenfeiten en tientaloverschrijding ($F(1,27) = 2.07$, ns). Kinderen, die de rekenfeiten beter beheersten, losten, in tegenstelling tot de verwachting, opgaven met tientaloverschrijding in vergelijking met opgaven zonder tientaloverschrijding niet sneller op dan kinderen met minder kennis van de rekenfeiten. Als hoofdeffect had tientaloverschrijding wel invloed op het oplossingstempo ($F(1,27) = 16.7$, $p < .001$). Opgaven met tientaloverschrijding werden trager ($M = 14.01$) opgelost dan opgaven zonder tientaloverschrijding ($M = 8.21$). Daarnaast kon een significant interactie-effect vastgesteld worden tussen de beheersing van de rekenfeiten en de telafstand ($F(1,27) = 7.83$, $p < .01$) (zie Figuur 2). Analyse van het simpele hoofdeffect van deze tweede orde interactie toonde aan dat kinderen, die over meer geautomatiseerde rekenfeiten beschikten opgaven met een kleine telafstand niet significant sneller oplosten ($M = 5.65$) dan opgaven met een grote telafstand ($M = 7.65$) ($F(1,10) = 3.08$, ns). Dit in tegenstelling tot kinderen, die de rekenfeiten niet geautomatiseerd hadden, die opgaven met een kleine telafstand wel significant sneller oplosten ($M = 12.55$) dan opgaven met een grote telafstand ($M = 18.89$) ($F(1,17) = 46.07$, $p < .01$).



Figuur 2 *Het tweede orde interactie-effect tussen beheersing van de rekenfeiten en telafstand.*

Juistheid van oplossen

In een zelfde variantie analytisch design met als afhankelijke variabele dit maal de juistheid van oplossen van de cijferopgaven kon geen significant interactie-effect vastgesteld worden voor de beheersing van de rekenfeiten en tientaloverschrijding. Er was wel sprake van een hoofdeffect van de beheersing van de rekenfeiten. Kinderen, die meer rekenfeiten geautomatiseerd beheersten, maakten significant meer opgaven goed dan kinderen die over minder geautomatiseerde rekenfeiten beschikten (($M = 26.72$ resp. $M = 20.8$); $F(1,27) = 9.28$, $p < .01$)).

Strategievaardigheid

De betrouwbaarheid van de scoring van de strategieën was voldoende hoog: Cohens kappa lag boven het vereiste niveau van .60 (range = .66 – 1.00; mediaan = .89). Vergelijking van de groepen met een betere en slechtere beheersing van de rekenfeiten liet geen significant verschil in geavanceerdheid van strate-

giegebruik zien ($t(16) = 0.5$, ns). Multivariate variantie analyse met de beheersing van de rekenfeiten als onafhankelijke variabele en de frequentie van het gebruik van een verkeerde strategie, tel-, splits-, handig reken- en rekenfeitenstrategie als afhankelijke variabelen, liet evenmin een verschil in strategiegebruik zien tussen de beide groepen ($F(5,12) = 0.25$, n.s.). Uit de univariate analyse kwam een verschil naar voren in het toepassen van telstrategieën door kinderen die beschikten over meer of minder geautomatiseerde rekenfeiten. Kinderen, die de rekenfeiten beter beheersten, maakten significant minder gebruik van telstrategieën dan kinderen, die de rekenfeiten niet beheersten ($F(1,16) = 4.75$, $p < .05$).

De percentages van voorkomen van de verschillende strategieën in de onderzoeksgroep, gebruikt in het tweede deel van het oorspronkelijke onderzoek ($n = 29$) waren: verkeerde strategie: 15.6%, telstrategie: 12.1%, splitsstrategie: 26.6%, handig rekenstrategie: 7.5%, rekenfeitenstrategie: 27.5% en overige strategieën: 8.8%. Verder werd nagegaan wat de invloed was van strategievaardigheid op de juistheid van oplossen. In een herhaalde metingen analyse ($2 \times 2 \times 2 \times 2 \times 2$) waarbij strategievaardigheid en beheersing van de rekenfeiten als tussen subjecten factoren en operatie, afstand en tientaloverschrijding als binnen subjecten factoren waren opgenomen, bleek de gezochte tweede orde interactie tussen strategievaardigheid en tientaloverschrijding niet significant ($F(1,14) = 2.37$, ns).

Foutenanalyse

De foutieve antwoorden op de cijfertoets werden gecategoriseerd. De meest gescoorde categorieën met percentage van voorkomen in de totale onderzoeksgroep ($n = 58$) waren:

- operatie-fout: (optellen i.p.v. aftrekken of omgekeerd): 19.6%
- inwissel-leenfout (niet of verkeerd inwisselen of lenen): 15.6%
- positie-fout (fout tegen het positiestelsel): 5.3%
- smaller-from-larger (omdraaien van de volgorde van de eenheden bij aftrekopgaven met tientaloverschrijding): 19.6%
- telfout (afwijking met + of − maximaal 1 van de oorspronkelijke uitkomst): 26%
- overige fouten: 13.9%

Uit de multivariate analyse met als onafhankelijke variabele beheersing van de rekenfeiten en als afhankelijke variabelen de eerder genoemde foutencategorieën, kwam een significant verschil in gemaakte fouten naar voren tussen kinderen met en zonder geautomatiseerde kennis van de rekenfeiten ($F(5,23) = 3.52$, $p < .05$). Bij nadere univariate analyse van deze gegevens bleek dat de twee groepen significant van elkaar verschilden in gemaakte telfouten ($F(1,27) = 9.54$, $p < .01$). Kinderen met meer kennis van de rekenfeiten maakten minder telfouten dan kinderen met minder kennis op dat terrein. Wat betreft de overige foutencategorieën was geen significant verschil aanwezig.

4 *Discussie*

Het besproken onderzoek biedt geen bevestiging voor de verwachting dat kinderen met meer rekenfeitenkennis bij het oplossen van cijferopgaven minder vertraagd worden door tientaloverschrijding dan kinderen zonder rekenfeitenkennis. Het veronderstelde faciliterende effect van meer geautomatiseerde rekenfeiten is bij tientaloverschrijding niet gevonden. De mate van automatisering bleek enkel terug te vinden in het sneller kunnen oplossen van opgaven met een grote telafstand.

Ook wat betreft de juistheid van oplossen van cijferopgaven kon geen significant tweede orde interactie-effect tussen de beheersing van rekenfeiten en tientaloverschrijding aangetoond worden. Kinderen, die niet beschikken over geautomatiseerde rekenfeiten, kunnen desondanks cijferopgaven met tientaloverschrijding relatief even goed oplossen als kinderen die wel een aantal rekenfeiten geautomatiseerd hebben. Er kon enkel een voor de hand liggend algemeen vaardigheidseffect vastgesteld worden: kinderen, die de rekenfeiten beter beheersen, maken ook meer opgaven goed.

Het beschikken over geautomatiseerde rekenfeiten is echter wel terug te vinden in het strategiegebruik. Kinderen, die meer kennis hebben van de rekenfeiten, tellen minder dan kinderen, die weinig kennis van de rekenfeiten hebben. Er kon een relatie vastgesteld worden tussen de vaardigheid in het oplossen van rekenfeiten en het strategiegebruik. Zover hier nagegaan kon worden, werkt deze vaardigheid slechts heel globaal via het strategie-

gebruik door naar de juistheid van oplossen. Kinderen, die de rekenfeiten gedeeltelijk geautomatiseerd hebben, maken met name minder telfouten in cijferopgaven dan kinderen waarbij de rekenfeitenkennis ontbreekt. De afname van het aantal telfouten en de vermindering van het gebruik van telstrategieën loopt wellicht parallel. Dit effect lijkt tot uitdrukking te komen in de juistheid van oplossen bij alle opgaventypen. Hoewel niet duidelijk is of het aantal telfouten ook daadwerkelijk afneemt doordat kinderen minder telstrategieën toepassen.

Er lijkt echter een specifiek tempo-effect te zijn bij opgaven met een grote telafstand. Voor kinderen, die over een aantal geautomatiseerde rekenfeiten beschikken, lijkt de telafstand geen differentiële invloed te hebben: ze lossen de opgaven met een kleine of grote telafstand even snel op doordat ze het antwoord mogelijk direct uit het geheugen kunnen ophalen of in ieder geval niet tellend hoeven te achterhalen. Kinderen, die voor hun oplossingsmethode aangewezen zijn op tellen, hebben wel aanzienlijk meer tijd nodig voor het oplossen van opgaven met een grote dan met een kleine afstand. Dit resultaat zou er op kunnen wijzen dat kennis van de rekenfeiten ook gebruikt wordt bij het oplossen van cijferopgaven. Bij deze bevinding dient echter aangetekend te worden dat de gehanteerde definitie van 'telafstand' wellicht niet adequaat is. Verondersteld werd dat zwakke kinderen het tweede getal uit de opgave geheel tellend bij het eerste getal zouden op- of aftellen. Dit geheel uittellen van de uitkomst komt veelvuldig voor bij horizontaal afgebeelde opgaven (Andrea, 1988). Uit, niet systematische, observaties van het oplossingsgedrag van kinderen in onderhavig onderzoek bleek echter dat zij bij verticale somnotatie voornamelijk kolomsgewijs te werk gingen. Dus ook bij de toepassing van een telprocedure werd per kolom door- of teruggeteld. Hiermee wordt het gehanteerde afstandsconcept enigszins ondergraven. Het enige relevante verschil tussen opgaven met een grote en kleine telafstand lijkt niet de absolute grootte van het getal te zijn, maar het aantal cijfers waaruit het tweede getal bestaat. Bij een kleine afstand hoeft slechts een ééncijferig getal opgeteld of afgetrokken te worden, terwijl bij een grote afstand een keer extra opgeteld of afgetrokken moet worden. Hiermee zou het tweede orde interactie-effect tussen de vaardigheid in het oplossen van rekenfeiten en afstand verstaan kunnen worden als zouden kinderen die geen geautomatiseerde rekenfeiten tot hun beschikking hebben door het één keer meer moeten optellen of aftrekken relatief meer vertraagd worden dan kinderen die over een aantal geautomatiseerde rekenfeiten beschikken. Aangezien het tellend toevoegen of aftrekken van een getal tussen 0 en 10 gemiddeld meer tijd kost dan het geautomatiseerd optellen of aftrekken van zo'n getal, leidt het tweemaal toepassen van deze telprocedures tot een verschil in oplossingstijd dat groter is dan het verschil in oplossingstijd bij tweemaal inzetten van geautomatiseerde rekenfeiten. Of er al dan niet ingewisseld moet worden, schijnt voor beide vaardigheidsgroepen hetzelfde vertragende effect te hebben.

Uit dit onderzoek wordt niet duidelijk welke gevolgen de mate van automatisering van de rekenfeiten, naast de teruggang in telstrategieën, heeft voor de gebruikte oplossingsprocedures. Het onderzoek biedt enerzijds door het ontbreken van duidelijke samenhang tussen beheersing van rekenfeiten en het gebruik van een rekenfeitenstrategie ondersteuning voor de studie van Resnick en Omanson (1986) waarin ook aangetoond werd dat de aanwezigheid van kennis van een andere orde, nl. conceptuele kennis, geen garantie vormde voor de toepassing van deze kennis in de oplossingsstrategie. Anderzijds blijkt er toch invloed te zijn van de kennis van rekenfeiten op het gebruik van telstrategieën en het aantal telfouten dat gemaakt wordt.

Evenals in het onderzoek van Russel en Ginsburg (1984) kon een verband vastgesteld worden tussen slechtere rekenprestaties (met betrekking tot het aantal goed gemaakte opgaven en het aantal telfouten) en een minder geautomatiseerd zijn van de rekenfeiten. Hiermee is echter niet gezegd dat onvoldoende kennis van de rekenfeiten de oorzaak is van rekenproblemen. Slechte rekenprestaties (m.n. inwisselfouten) kunnen ook toegeschreven worden aan gebrekkige strategievaardigheid. Kennis van de rekenfeiten lijkt vooral van belang voor nauwkeurig rekenen i.c. minder telfouten.

Op basis van dit onderzoek kan het specifieke effect van het geautomatiseerd zijn van de rekenfeiten voor tientaloverschrijding niet aangetoond worden. Het effect zou wellicht

wel tot uitdrukking gekomen zijn als meer variatie in de onderzoekspopulatie aangebracht was. In de huidige onderzoeksgroep zaten slechts enkele kinderen die de cijferopgaven foutloos beheersten. Hierdoor is de groep misschien te homogeen van samenstelling geweest waardoor specifieke effecten niet aangetoond konden worden.

Een tweede verklaring voor dit tegenvallend resultaat wordt gevormd door het feit dat in dit onderzoek de opgaven in verticale somvorm werden aangeboden. Hierdoor wordt eerder een cijferende dan hoofdrekenende aanpak uitgelokt. Doordat bij cijferend rekenen, in tegenstelling tot hoofdrekenen, de tussenuitkomsten opgeschreven worden, staat het werkgeheugen minder onder druk. Het is mogelijk dat het geautomatiseerd beheersen van de rekenfeiten voor het cijferend rekenen minder belangrijk is dan voor het hoofdrekenen. Het ontbreken van de rekenfeiten staat waarschijnlijk het ontstaan van de kennis nodig bij het leren inwisselen en lenen niet in de weg. Hiermee willen wij niet zeggen dat oefenen van de rekenfeiten uit het rekencurriculum geschrapt kan worden. Rekenfeiten zijn van invloed op de nauwkeurigheid van rekenen. Daarnaast lijkt het zeer aannemelijk dat het vlot kunnen oplossen van de rekenfeiten de efficiëntie bij de verschillende rekenoperaties (ook delen en vermenigvuldigen) vergroot. Uit het onderzoek blijkt dat ook kinderen uit het speciaal onderwijs als ze de rekenfeiten enigszins beheersen, profijt hebben van deze kennis bij het correct oplossen van cijferopgaven. Het dient echter aanbeveling om in het rekencurriculum de rekenfeiten niet geïsoleerd aan te bieden, maar juist in combinatie met het toepassen van de rekenfeiten in de oplossingsstrategie.

Literatuur

Anderson, J. R., Acquisition of cognitive skill. *Psychological Review*, 1982, *89*, 369-406.

Andrea, M. C. *De bijdrage van de vaardigheid in het splitsen van getallen tot en met 10 aan de prestaties van het cijferend rekenen onder de 100*. Nijmegen: doctoraalscriptie orthopedagogiek, 1988.

Ashcraft, M. H., The development of mental arithmetic: a chronometric approach. *Developmental Review*, 1982, *2*, 213-236.

Ashcraft, M. H. & J. Battaglia, Cognitive arithmetic: evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology*, 1978, *4*, 527-538.

Baltussen, M. W. J. & E. C. D. M. van Lieshout, De plaats van het splitsen in het cijferend optellen en aftrekken tot 100. *Tijdschrift voor Onderwijswetenschappen*, 1989, *20*, 52-63.

Cauly, K. M., Construction of logical knowledge: study of borrowing in subtraction. *Journal of Educational Psychology*, 1988, *80*, 202-205.

Eysenck, M. W. *A handbook of cognitive psychology*, London: Erlbaum, 1984.

Fuson, K. C., An analysis of the couting-on solution procedure in addition. In: T. P. Carpenter; J. M. Moser & T. A. Romberg (Eds.), *Addition and subtraction; a cognitive perspective*, Hillsdale, NJ: Erlbaum, 1982.

Greeno, J. G., Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist*, 1978, *12*, 262-283.

Groen, G. J. & J. M. Parkman, A chronometric analysis of simple addition. *Psychological Review*, 1972, *79*, 329-343.

Hamann, M. S. & M. H. Ashcraft, Simple and complex mental addition across development. *Journal of Experimental Child Psychology*, 1985, *40*, 49-72.

Laberge, D. & S. J. Samuels, Towards a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, *6*, 293-323.

Lucas, P. J. & L. C. van der Gaag, *Principes van expertsystemen*. Schoonhoven: Academic Service, 1988.

Payne, S. J., Methods and mental models in theories of cognitive skills. In: J. Self (Ed.), *Artificial intelligence and human learning*. London: Chapman and Hall, 1988.

Perfetti, C. A. & A. M. Lesgold, Coding and comprehension in skilled reading and implications for reading instruction. In: L. B. Resnick & P. A. Weaver (Eds). *Theory and practice of early reading*, Vol 1. Hillsdale, NJ: Erlbaum, 1979.

Resnick, L. B. & W. W. Ford, *The psychology of mathematics for instruction*. London: Erlbaum, 1984.

Resnick, L.B. & S.F. Omanson, Learning to understand arithmetic. In: R. Glaser (Ed.), *Instructional Psychology*, Vol. 3. Hillsdale, NJ: Erlbaum, 1987.

Russel, R.L. & H.P. Ginsburg, Cognitive analysis of children's mathematics difficulties. *Cognition and Instruction*, 1984, *1*, 217-244.

Steinberg, R.M., Instruction on derived facts strategies in addition and subtraction. *Journal for Research in Mathematics Education*, 1985, *16*, 337-355.

Usnick, V. & J.M. Engelhardt, Basic fact, numeration concepts and the learning of the standard multidigit addition algorithm. *Focus on Learning Problems in Mathematics*, 1988, *10*, 1-14.

Wolters, M., Die twee streepjes boven elkaar, is dat een "is"?. *Willem Bartjens*, 1984/1985, *4*, 216-221.

Curricula vitae

M. W. J. Baltussen (1962) studeerde orthopedagogiek te Nijmegen. Tijdens haar studie was ze korte tijd werkzaam aan een onderwijsbegeleidingsdienst. Sinds haar afstuderen in 1986 werkt ze aan het Instituut voor Orthopedagogiek te Nijmegen, al waar ze onderzoek doet op het gebied van de theorie en behandeling van rekenproblemen.

E. C. D. M. van Lieshout: zie *Pedagogische Studiën*, 1989, *66*, p. 255.

Adres: Instituut voor Orthopedagogiek, Katholieke Universiteit Nijmegen, Postbus 9103, 6500 HD Nijmegen.

Manuscript aanvaard 29-3-'90

Summary

Baltussen, M.W.J. & E.C.D.M. van Lieshout. 'Number facts in paper-and-pencil addition and subtraction.' *Pedagogische Studiën*, 1990, *67*, 306-316.

Knowing number facts by heart is assumed to be important for paper-and-pencil addition and subtraction. Thirty-nine children with learning disabilities and 19 mildly retarded children were examined. They were administered a number facts test, an addition and subtraction test and a thinking aloud test. ANOVA showed that children who were more proficient in solving number facts compared to children less gifted in that area, solved sums faster. Contrary to what was expected the difference was not larger for sums with than for sums without borrowing and carrying. The difference was larger for sums with double digit addends or minuends than for sums with single digit addends or minuends. A relation could also be established between the proficiency in solving number facts and the correctness of solving the sums. The more proficient children less frequently computed the answers by counting and made fewer counting errors than children who did not know any number facts by heart.

Bijlage 1

De cijfertoets

De opgaven staan horizontaal afgebeeld. De kinderen kregen ze echter in verticale vorm voorgelegd. Ook de volgorde was anders.

optellen zonder tiental-overschrijding	optellen met tiental-overschrijding	aftrekken zonder tiental-overschrijding	aftrekken met tiental-overschrijding
42 + 33 =	48 + 35 =	75 – 32 =	72 – 35 =
32 + 54 =	38 + 56 =	96 – 52 =	94 – 56 =
61 + 5 =	69 + 6 =	76 – 1 =	71 – 6 =
24 + 63 =	26 + 67 =	78 – 45 =	93 – 67 =
32 + 5 =	48 + 7 =	47 – 2 =	42 – 7 =
25 + 43 =	25 + 48 =	48 – 2 =	75 – 48 =
37 + 2 =	34 + 8 =	68 – 31 =	42 – 8 =
21 + 37 =	29 + 38 =	97 – 63 =	61 – 38 =

Bijlage 2

De strategie-toets

De opgaven staan hier horizontaal afgebeeld. De kinderen kregen ze echter in verticale vorm voorgelegd.

48 + 35 =	29 + 38 =
71 – 6 =	34 + 8 =
93 – 67 =	72 – 35 =
26 + 67 =	69 + 6 =
42 – 8 =	61 – 38 =

Brugklassers met discordante lees- en spelprestaties

W. H. J. VAN BON en F. G. M. BOUWENS
Instituut voor Orthopedagogiek
Katholieke Universiteit Nijmegen

Samenvatting

Lees- en spelprestaties hangen niet perfect met elkaar samen. Het blijkt dan ook mogelijk brugklasleerlingen te vinden die beter spellen dan ze lezen en brugklassers bij wie het omgekeerde het geval is. Om na te gaan of deze discordantie van lezen en spellen berust op strategieverschil, is onderzocht hoe goede-spellers-zwakke-lezers (GSZL's) en zwakke-spellers-goede-lezers (ZSGL's) pseudowoorden, regelmatige en onregelmatige woorden spellen (nauwkeurigheid) en lezen (nauwkeurigheid en latentietijd).

De uitkomsten suggereren dat de GSZL's lezen met een fonologische strategie en op grond van volledige woordinformatie waardoor ze de orthografische kennis opdoen die hen in staat stelt onregelmatige woorden beter te schrijven dan de ZSGL's. Dat de ZSGL's niettemin op een cloze-toets de betere lezers bleken te zijn zou erop wijzen dat zij beter gebruik weten te maken van contextuele informatie, terwijl de GSZL's daarin belemmerd worden door de capaciteits-eisen die hun woordleesstrategie stelt.

1 *Inleiding*

Lezen en schrijven worden vaak als twee kanten van dezelfde medaille beschouwd. De vaardigheden zouden nauw met elkaar verweven zijn en vrijwel tegelijkertijd worden aangeleerd. Wie kan lezen kan spellen; wie kan spellen kan lezen. Argumenten zijn gemeenschappelijkheid van deelprocessen, gebruik in het een (spellen) van kennis die afkomstig is uit het ander (lezen), verwevenheid in het onderwijs en gevoeligheid voor dezelfde belemmerende of bevorderende factoren. Als die opvatting juist is dan zou, binnen de grenzen die gesteld worden door de betrouwbaarheid van de meetinstrumenten, het verband tussen lees- en spelprestaties perfect moeten zijn.

Uit de onderzoeksliteratuur (bijv. Boland & Mommers, 1986; Van Bon, 1984; Malmquist, 1958; Shanahan, 1984) blijkt echter dat het verband niet perfect is. Bryant en Bradley (1980) deden de verrassende constatering dat kinderen woorden soms goed schrijven ondanks dat ze ze fout lezen. Uit onderzoek van Frith (1980) is bekend dat er kinderen zijn die goede leesvaardigheid combineren met zwakke spellingvaardigheid. Deze discordantie duidt zij aan met de term 'unexpected poor spelling'.

Ook de voorbereidende fase van het onderzoek waarvan hier verslag wordt gedaan biedt empirische ondersteuning voor het bestaan van discordantie tussen leesvaardigheid en schrijfvaardigheid (Bouwens, 1988). In het vooronderzoek, en in het aansluitende onderzoek waarvan in dit artikel verslag wordt gedaan, is lezen in eerste instantie opgevat als begrijpend lezen, omdat dat de uiteindelijk te verklaren vaardigheid is (Perfetti, 1985). Leesvaardigheid is daarom bepaald met een integratieve of globale maat waarin ook tekstbegrip tot uitdrukking komt. Een cloze-toets is daarvoor een gemakkelijk te contrueren en valide middel (Rost, 1989). Hij werd opgezet volgens een gebruikelijke procedure. Elk zevende woord in een niet te moeilijke tekst (uit "Pjotr" van Terlouw) was vervangen door een liggende streep. De taak voor de proefpersonen was de weggelaten woorden in te vullen. Tot globale maat voor de spellingvaardigheid diende een auditief woorddictee waarin een grote verscheidenheid aan spellingproblemen aan de orde komt (Geerling & Geurts, 1986). Deze toetsten zijn voorgelegd aan 266 brugklassers met een gemiddelde leeftijd van twaalf jaar en negen maanden. De leerlingen waren allen afkomstig van een school voor Voortgezet Onderwijs waar brugklasleerlingen zijn verdeeld over LBO/MAVO-brugklassen en MAVO/HAVO/VWO-brugklassen.

De correlatie tussen het aantal goed op de cloze-toets en dat op de spellingtoets was 0.54.

De globale maat voor leesvaardigheid en die voor spellen hangen dus niet perfect met elkaar samen[1]. Hoewel er een zeker verband bestaat, verklaart hun correlatie minder dan 1/3 van de variantie. Goede leesvaardigheid gaat soms gepaard met zwakke spellingvaardigheid; ook het omgekeerde blijkt voor te komen. De constatering dat lees- en spelvaardigheid niet perfect met elkaar correleren, komt overeen met de bevindingen van, bijvoorbeeld, Aarnoutse, Mommers, Smits en Van Leeuwe (1986).

Het lijkt dus mogelijk kinderen te vinden die goed zijn in het een en zwak in het ander: zwakke spellers die goed lezen (ZSGL's) en goede spellers die zwak lezen (GSZL's). Op basis van hun cloze- en spelscores in het onderzoek van Bouwens (1988) werden leerlingen geselecteerd die tot een van deze twee groepen konden worden gerekend. Als zwak resp. goed werden de leerlingen beschouwd die behoorden tot de 40% laagst resp. hoogst scorenden van het betreffende brugklasniveau (LBO/MAVO of MAVO/HAVO/VWO). Deze indeling leverde 42 leerlingen op die tot een discordante groep behoren en dus goed zijn in het een maar zwak in het ander: 21 ZSGL's en 21 GSZL's.

De *centrale onderzoeksvraag* betreft deze twee groepen met discordante prestaties: Hoe onderscheiden leerlingen die zwak spellen en goed lezen (ZSGL) zich van leerlingen die goed spellen en zwak lezen (GSZL) in de aard van hun lezen en hun spellen? Getracht wordt het lezen en spellen van deze groepen leerlingen te karakteriseren in termen van twee strategieën die men zou kunnen volgen bij het uitvoeren van beide taken. De ene strategie bestaat uit toepassing van het fonologisch principe, de andere berust op het gebruik van woordspecifieke orthografische informatie (zie bijv. Ellis, 1984). Toepassing van het fonologische principe bij het lezen, de fonologische strategie, stelt men zich voor als de omzetting van de opeenvolgende letters of lettergroepen in een fonologische code, gevolgd door betekenisidentificatie en/of spraakproduktie. Bij het spellen zou een fonologische strategie of werkwijze bestaan uit opdeling van het gesproken woord in segmenten die via correspondentieregels worden omgezet in letters of lettergroepen. Deze werkwijze maakt het mogelijk woorden die men nog niet eerder gezien heeft, bijvoorbeeld omdat ze niet bestaan (pseudowoorden), of waarvan men zich de schrijfwijze niet meer herinnert, te lezen en te schrijven.

De andere werkwijze, de orthografische strategie, maakt gebruik van woordspecifieke informatie. Dat is nodig om onregelmatige woorden (zoals CACAO, JAM) efficiënt te herkennen en correct te kunnen schrijven, om bij het het lezen de juiste betekenis van homofonen (MEID, MIJT) te bepalen en bij het spellen, gegeven een betekenis, te kiezen tussen homofone spellingen.

Een verklaring voor het uit de pas lopen van lees- en spelvaardigheid is dat bij het spellen een fonologische strategie zou overheersen terwijl bij het lezen gebruik van woordspecifieke kennis de overhand heeft (Barron, 1980; Bryant & Bradley 1980; Frith, 1980). Er zijn inderdaad aanwijzingen dat de bekwaamheid in het toepassen van de ene strategie niet noodzakelijk of volledig samengaat met die in het toepassen van de andere (cfr. Van Aarle & Vollebergh, 1986). Eén bron van evidentie vormen personen die ooit volleerde lezers en schrijvers waren en bijvoorbeeld ten gevolge van een hersenbloeding niet meer in staat zijn tot het lezen of schrijven van bepaalde categorieën woorden, bijvoorbeeld pseudowoorden. Dat kan worden opgevat als aanwijzing dat de betreffende werkwijze, in dit voorbeeld de fonologische, een relatief zelfstandig onderdeel vormde van de cognitieve competentie van deze personen (Ellis, 1984). Een andere bron van evidentie vormt correlationeel onderzoek. Dat van Treiman (1984) is een voorbeeld. Zij constateerde dat de vaardigheid van kinderen in het spellen van pseudowoorden hoger correleerde met hun vaardigheid in het spellen van regelmatige, d.i. volgens niet-woordspecifieke spellingvoorschriften geschreven woorden dan met hun vaardigheid in het schrijven van onregelmatige woorden. De correlatie tussen regelmatige en onregelmatige woorden lag tussen die van pseudowoorden met regelmatige resp. onregelmatige woorden. Voor lezen vond zij een zelfde correlatiepatroon. Bovendien bleek dat degenen die de fonologische lees- en spellingregels lijken te volgen meer fouten maken die 'fonetisch' correct zijn. Ook dit wijst erop dat bekwaamheid in het toepassen van de fonologische strategie en in het gebruik van woordspecifieke associaties niet volledig met elkaar samenhangen. Treiman concludeerde bovendien dat lees- en

spelstrategie in de regel overeenkomen.

Het beeld dat Frith (1985) schetst van de samenhang tussen lees- en spelstrategie is echter complexer. Ze beschrijft de normale ontwikkeling als een verschuiving van lees- en spelstrategie. Na een aanvankelijk logografisch stadium, waarin saillante grafische kenmerken bij lezen en spellen de hoofdrol spelen, zou het spellen een alfabetische (d.i. fonologische) fase ingaan waarin de processen voornamelijk berusten op foneem-grafeem-correspondentieregels. Dit zou vergelijkbare alfabetische processen bij het lezen op gang brengen die daar echter spoedig plaats maken voor orthografische processen die berusten op kennis van de woordstructuur. Pas in de volgende fase dient deze orthografische informatie als basis voor het spellen, dat tot dan toe overwegend alfabetisch van aard is gebleven. De normale ontwikkeling zou dus gekarakteriseerd worden door het soms samengaan, soms uit de pas lopen van lees- en spellingstrategie. Individuele verschillen in de overeenstemming tussen lees- en spellingstrategie kunnen in dit kader geïnterpreteerd worden als gevolg van ontwikkelingsvertraging of -stilstand. In dat verband vat Frith het samengaan van zwak spellen en goed lezen op als het blijven voortbestaan van een fonologische manier van spellen naast een orthografische leesstrategie. Dat uit zich in het correct lezen van onregelmatige woorden en homofonen (THEIR, THERE), maar moeizaam en vaak foutief spellen op basis van klank-letter-koppelingsregels. Frith verklaart dat uit een geringe mate van gedetailleerdheid van de interne (lexicale) representatie van de orthografische eenheden. Bij goede lezers die zwak spellen zouden deze representaties voldoende zijn voor woordherkenning, maar te imprecies voor spellen, waardoor het spellen op de minder geavanceerde fonologische manier moet gebeuren. Personen die spellen in overeenstemming met hun leesvaardigheid zouden preciezer representaties opdoen bij het lezen door gelijke aandacht te geven aan alle letters in het woord, hetgeen meer aandacht is dan voor lezen noodzakelijk is. "It is conceivable that individual differences exist in terms of willingness/capacity to adopt such a wastefully inelegant strategy, and this would provide an explanation for arrest at this point in the sequence" (Frith, 1985, blz. 321). Die verkwistende manier van lezen, "reading by full cues", karakteriseert zij elders (Frith, 1980, blz. 506) ook als "reading by ear"; een manier van lezen die vooral tot uitdrukking zou komen bij het lezen van nonsensewoorden, waarbij het aankomt op toepassing van grafeem-foneem-correspondenties en op aandacht voor de klank. Het lijkt er dus op dat volgens Frith de preciezer orthografische representaties tot stand komen met behulp van een fonologische strategie.

Het is opmerkelijk dat Frith geen aandacht besteedt aan de tegenovergestelde discordantie, onverwacht slecht lezen. Haar redenering voortzettend ligt het immers voor de hand dat er personen te vinden zullen zijn die lezen op zo'n onelegant nauwkeurige manier waarmee ze meer orthografische woordinformatie opdoen dan noodzakelijk is voor woordherkenning en de precisie bereiken die noodzakelijk is voor correct spellen, maar daarbij aan de grens van hun verwerkingscapaciteit raken zodat leesprocessen voorbij het woordniveau belemmerd worden. Personen, met andere woorden, die wel goed zijn in technisch lezen van afzonderlijke woorden maar niet in vormen van lezen waarbij contextgebruik een belangrijke rol speelt, zoals de cloze-taak.

Van de discordante lezers/spellers zullen de ZSGL's dan gekenmerkt zijn door gebrekkige identificatie van pseudowoorden en goede herkenning maar relatief zwakke spelling van onregelmatige woorden; de GSZL's, die zich negatief onderscheidden in een leestaak waarin gebruik van contextuele informatie van belang is, door relatief goed lezen van pseudowoorden en goed spellen van onregelmatige woorden.

Deze voorspelling wordt getoetst door de snelheid resp. de nauwkeurigheid te vergelijken waarmee beide groepen leerlingen onregelmatige en pseudowoorden lezen en de nauwkeurigheid waarmee ze deze woordtypen spellen. Ter verkenning van het effect van woordbekendheid zijn ook regelmatige woorden in het onderzoek betrokken.

2 *Methode*

2.1 *Proefpersonen*

De leerlingen die in de eerste fase van het onderzoek als ZSGL's of GSZL's waren geïndentificeerd, fungeerden als proefpersonen in dit deel van het onderzoek. Er waren 21 leer-

lingen van elk type, waarvan 9 uit brugklassen van LBO/MAVO-niveau, 12 uit brugklassen van MAVO/HAVO/VWO-niveau. Onder de ZSGL's waren 11 meisjes, 5 onder de GSZL's. De gemiddelde leeftijd van de ZSGL's en de GSZL's was resp. 13 jaar, 0 maanden en 12 jaar, 11 maanden.

2.2 *Toetsen*

Lezen. Regelmatige, onregelmatige en pseudowoorden werden één voor één aangeboden in een taak waarbij de leerlingen de opdracht hadden elk woord snel en nauwkeurig te benoemen, de hardop-leestaak (HLT). Er waren 20 woorden van elk type. Woorden heetten regelmatig als ze geschreven zijn volgens de hoofdregels in de systematiek van de Nederlandse spelling. Voor meerlettergrepige woorden houdt dat in dat samenstellingen 'klankzuiver' geschreven zijn en dat niet-samengestelde woorden geschreven zijn volgens de regels voor de open en de gesloten lettergreep en overigens klankzuiver zijn. Voorbeelden zijn VERWORDING, PRINSES, VRIESPUNT. De gebruikte onregelmatige woorden wijken af van deze hoofdregels in de spellingsystematiek: CHIMPANSEE, YOGHURT, CIRCUS. De pseudowoorden werden door klinkervervanging afgeleid uit regelmatige woorden: KONDELUUR, FAGOOR, WOEPENBREDER. De woorden van elk type waren gemiddeld even lang in aantal letters (7.7), aantal lettergrepen (2.8) en kwamen overeen in klinker/medeklinkerratio (.74). Voor de keuze van regelmatige en onregelmatige woorden die naar gebruiksfrequentie en moeilijkheidsgraad geschikt zijn voor ons doel werd gebruik gemaakt van de resultaten van een onderzoek waarin spellingtoetsen zijn ontwikkeld voor de verschillende leerjaren van LBO, MAVO en HAVO (Geerling & Geurts, 1986).

De woorden werden in kleine letter aangeboden in het midden van een beeldscherm. De latentietijd tussen de presentatie van het woord en het antwoord van de proefpersoon werd middels een 'voice key' geregistreerd door een computer. Randapparatuur zorgde ervoor dat de proefpersonen zelf het begin van de woordpresentatie konden bepalen en dat de proefleider lees- en procedure-fouten kon registreren. Een woord bleef in beeld totdat de proefpersoon zijn antwoord gegeven had.

De afname gebeurde individueel. Met 15 oefenwoorden werden de proefpersonen vertrouwd gemaakt met opstelling, procedure en pseudowoorden. Onmiddellijk daarna werden de 60 toetswoorden gelezen. De woordtypen wisselden elkaar daarbij af in een toevalsvolgorde die voor alle proefpersonen dezelfde was.

Spellen. Een woorddictee (WD) van 14 regelmatige, 26 onregelmatige en 14 pseudowoorden werd individueel afgenomen. Bestaande woorden waren geselecteerd met dezelfde criteria als die voor de HLT. Pseudowoorden werden door klinkervervanging afgeleid uit regelmatige woorden. De woordtypen waren wederom nagenoeg gelijk in gemiddeld aantal letters (7.9), aantal lettergrepen (2.5) en klinker/medeklinkerratio (.72). De woorden werden met een bandopname aangeboden. Elk woord werd daarbij eenmaal herhaald en duidelijk uitgesproken. De woordtypen wisselden elkaar weer af in een toevalsvolgorde die voor alle proefpersonen dezelfde was.

3 *Resultaten*

Tabel 1 geeft de gemiddelde scores van ZSGL's en GSZL's op de gebruikte toetsen met de bijbehorende standaarddeviaties.

Een variantie-analyse volgens de multivariate methode met het *aantal correct gelezen* woorden als afhankelijke variabele, leerlingtype als tussen-proefpersonen-factor en woordtype als binnen-proefpersonen-factor, laat een significante interactie zien ($F(2,39) = 9.66$; $p<0.01$). De accuratesse waarmee verschillende woordtypen gelezen worden, onderscheidt tussen de typen leerlingen (figuur 1). De aansluitende univariate toetsing laat zien dat het verschil tussen de pseudo- en de regelmatige woorden significant anders is voor beide groepen leerlingen ($F(1,40) = 19.24$; $p<0.01$). Het verschil tussen pseudo- en onregelmatige woorden is bij de GSZL's kleiner dan bij de ZSGL's. Het verschil tussen regelmatige en de onregelmatige woorden verschilt niet voor de beide groepen leerlingen ($F(1,40) = 0.04$; $p = 0.84$). Het hoofdeffect van woordtype is significant ($F(2,39) = 61.12$; $p<0.01$): de leerlingen van beide groepen lezen regelmatige woorden nauwkeuriger dan onregelmatige ($F(1,40) =$

Tabel 1 *Hardop-leestoets en woorddictee gemiddelden (standaarddeviaties)*

	Leerlingtype ZSGL (n = 21)	GSZL (n = 21)
Hardop leestoets		
Aantal goed		
Regelm. woorden	18.48 (0.93)	18.91 (1.09)
Onregelm. woorden	16.76 (1.04)	7.0 (1.30)
Pseudo-woorden	13.24 (2.41)	16.48 (2.36)
Latentietijd		
Regelm. woorden	754 (143)	738 (95)
Onregelm. woorden	797 (166)	782 (98)
Pseudo-woorden	997 (223)	1015 (159)
Woorddictee		
Percentage goed		
Regelm. woorden	91.16 (6.35)	94.22 (5.35)
Onregelm. woorden	40.29 (8.39)	64.65 (10.37)
Pseudo-woorden	73.47 (11.55)	77.55 (10.66)
Foutentype onregelm. woorden (aantal)		
Fonetische	10.43 (2.48)	5.81 (2.36)
Niet-fonetische	5.10 (2.17)	3.38 (2.11)

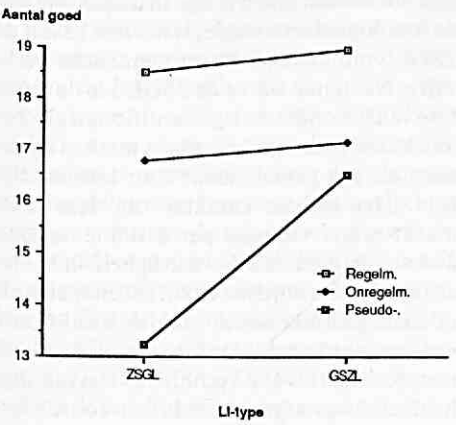
123.82; $p<0.01$) en die weer beter dan pseudowoorden ($F(1,40)=39.14$; $p<0.01$). Ook het hoofdeffect van leerlingtype is significant ($F(1,40)=13.93$; $p=0.01$). Uit de significante interactie moeten we echter concluderen dat het verschil tussen de leerlingtypen varieert met het woordtype. Het verschil is het grootst voor de pseudowoorden. Voor de regelmatige woorden zou het beperkt kunnen zijn door een plafond-effect.

De GSZL's lijken regelmatige woorden en onregelmatige woorden sneller te lezen dan de ZSGL's. Bij de pseudo-woorden lijkt het omgekeerde het geval te zijn (Figuur 2). Uit een multivariate variantie-analyse met *benoemsnelheid* als afhankelijke variabele, leerlingtype en woordtype als design-factoren blijkt echter geen significante ineractie van leerlingen woordtype ($F(2,39)=0.55$; $p=0.58$): mocht er een verschil zijn in de snelheid waarmee de onderscheiden woordtypen benoemd worden, dan is dat verschil voor beide groepen leerlingen hetzelfde. Een significant snelheidsverschil op de pseudowoorden had overigens een artefact kunnen zijn van het feit dat de benoemsnelheid alleen bepaald is over de correct gelezen woorden. De GSZL's lazen meer pseudowoorden goed, dus ook de moeilijkere die een langere latentietijd hebben: de correlatie tussen de moeilijkheidsgraad per woord (het percentage goede benoemingen) en zijn gemiddelde latentietijd was gemiddeld over de 60 woorden -.46.

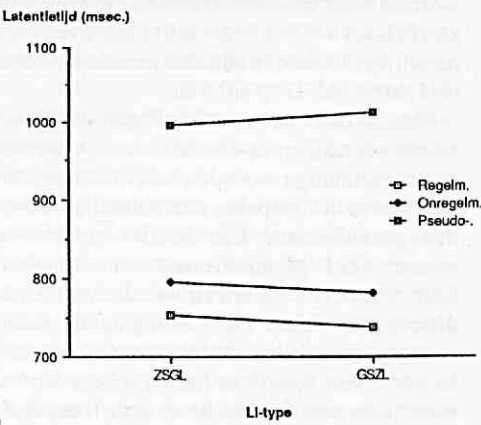
Er is een significant effect van woordtype ($F(2,39)=128.19$; $p<0.01$). Regelmatige woorden worden sneller gelezen dan onregelmatige ($F(1,40)=252.32$; $p<0.01$). Onregelmatige woorden worden echter sneller benoemd dan pseudo-woorden ($F(1,40)=180.08$; $p<0.01$).

Het snelheidsverschil tussen de ene groep leerlingen en de andere is niet significant ($F(1,40)=0.01$; $p=0.92$).

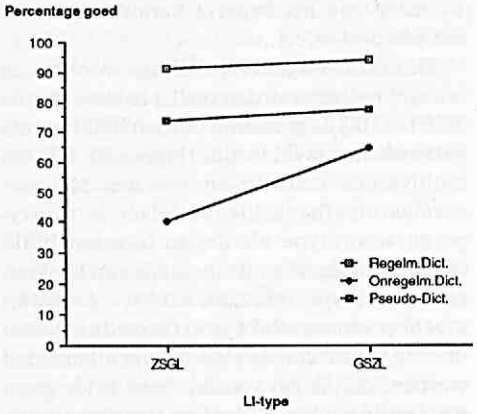
Een multivariate variantie-analyse met de *spellingscore* als afhankelijke variabele laat



Figuur 1 *Hardop-leestoets, aantal goed gelezen woorden.*



Figuur 2 *Hardop-leestoets, latentietijd.*



Figuur 3 *Woorddictee, percentage goed geschreven woorden.*

een significante interactie van leerlingtype en woordtype zien ($F(2,39) = 21.36$; $p < 0.01$). Zowel het verschil tussen de regelmatige en de onregelmatige woorden als dat tussen de onregelmatige en de pseudowoorden is kennelijk voor beide groepen niet hetzelfde: $F(1,40) = 39.08$ resp. 27.79, $p < 0.01$. Het verschil tussen deze woordtypen is voor de GSZL's kleiner dan voor de ZSGL's (Figuur 3). Dat komt kennelijk omdat de GSZL's beter zijn in het spellen van onregelmatige woorden ($F(1,40) = 70.06 = p < 0.01$); het verschil met betrekking tot de beide andere woordtypen is niet significant. Het significante hoofdeffect van leerlingtype ($F(1,40) = 30.75$; $p < 0.01$) berust dus op de spelling van de onregelmatige woorden. Ook het verschil tussen de woordtypen is significant ($F(2,39) = 272.58$; $p < 0.01$). De regelmatige woorden worden beter geschreven dan de onregelmatige ($F(1,40) = 409.15$; $p < 0.01$) die overigens moeilijker blijken te zijn dan pseudowoorden ($F(1,40) = 143.47$; $p < 0.01$).

Omdat de twee typen leerlingen enkel van elkaar verschillen in de accuratesse waarmee ze onregelmatige woorden schrijven, zijn alleen de foutief gespelde onregelmatige woorden geanalyseerd. Een foutief geschreven woord werd geclassificeerd als 'fonetisch fout' wanneer de klanken van de standaarduitspraak volgens de spellingregels waren weergegeven, anders als 'niet-fonetisch fout'. In geval van meerdere fouten in een woord waaronder niet-fonetische werd de foute spelling ingedeeld bij de niet-fonetische. Leerlingen van beide typen maken meer fonetische dan niet-fonetische fouten, maar de ZSGL's overtreffen daarin de GSZL's: 72% vs. 61% ($F(1,40) = 4.78$, $p < 0.05$).

4 *Discussie*

ZSGL's en GSZL's verschillen in een aantal opzichten van elkaar. Bij de hardopleestaak overtreffen de GSZL's de ZSGL's in nauwkeurigheid als pseudowoorden gelezen moeten worden, dus in geval van woorden waarbij de fonologische werkwijze gebruikt moet worden. Er is geen accuratesseverschil in geval van regelmatige en onregelmatige woorden. De manier waarop ZSGL's lezen geeft hen kennelijk geen voordeel bij de identificatie van een van de onderscheiden woordtypen. Ook in snelheid overtreffen ze de GSZL's niet; de groepen onderscheiden zich niet van elkaar in de snelheid waarmee ze de drie woordtypen lezen. Met dezelfde snelheid behalen de GSZL's bij pseudowoorden een grotere nauwkeurigheid dan de ZSGL's, mogelijk zelfs met grotere snelheid vanwege de eerder gemelde samenhang tussen woordmoeilijkheid en latentietijd. Het lijkt er dus op dat de GSZL's vaardiger, d.w.z. preciezer, maar misschien ook sneller zijn in het toepassen van een fonologische werkwijze. De ZSGL's munten niet uit in het lezen van de andere woordtypen, de regelmatige en onregelmatige woorden; zij zijn daarin niet preciezer en ook niet sneller. Zij beschikken dus niet over efficiënter routines voor de identificatie van orthografische patronen of over een uitgebreider 'leeslexicon'. Terwijl de GSZL's relatief bekwaam lijken te zijn in toepassing van de fonologische strategie, is het niet zo dat de ZSGL's uitblinken in de orthografische werkwijze. Niettemin waren de ZSGL's in de eerste fase van het onderzoek geïdentificeerd als 'betere lezers'. Nu was dat gedaan met een clozetoets als een globale maat voor leesvaardigheid. Het globale karakter van deze toets maakt echter welhaast per definitie dat ook andere factoren dan de vaardigheid in het lezen van losse woorden een rol zullen spelen als de toets gebruikt wordt voor de kwalificatie van leesvaardigheid. De lage correlatie (0.20) met een toets voor technische leesvaardigheid[2] die was afgenomen in het vooronderzoek (Bouwens, 1988) onderstreept dat. Het feit dat de GSZL's op die toets hoger scoorden

dan de ZSGL's ($F(1,38)=24.29$; $p<0.01$) suggereert zelfs dat de betere leesvaardigheid van de ZSGL's op de cloze-toets aan andere factoren dan technische leesvaardigheid of decodeervaardigheid moet worden toegeschreven. Tot die andere factoren behoort het gebruik van informatie binnen zinnen en, in mindere mate (Shanahan, Kamil & Tobin, 1982; maar zie Henk, 1982), over zinsgrenzen heen om daaruit af te leiden hoe een open plaats in de zin moet worden opgevuld. Misschien is het wel dit contextgebruik waarin de ZSGL's beter zijn; hebben zij daarom niet zoveel orthografische informatie nodig om een woord in een zin te identificeren en gebruiken ze ook minder orthografische informatie. Zij zouden dan bij het lezen van teksten gebruik maken van informatie van hoger (syntactisch en semantisch) niveau en woorden al herkennen op grond van beperkte gegevens, zoals de eerste letters en een indruk van de woordlengte. Volgens Frith (1980; zie boven) zou dit "reading by partial cues" een economische en succesvolle werkwijze zijn in de meeste leessituaties van alledag. Maar zij zal slechts weinig gelegenheid bieden om zich spellingkennis eigen te maken zowel van de diepere spellingsystematiek als van woordspecifieke letterpatronen. Wie zo leest zal de spellinginformatie missen om afwijkende woorden correct te kunnen schrijven en toevlucht moeten nemen tot een fonologische manier van spellen. De slechte spellers die goed lazen in het onderzoek van Frith (1980) gedroegen zich inderdaad in overeenstemming met dit patroon. Ook de ZSGL's in ons onderzoek gedragen zich of zij woordspecifieke informatie niet hebben 'opgepikt' en bij onbekendheid met de schrijfwijze de in de ontwikkeling vroege fonologische werkwijze hanteren: woorden waarop de hoofdregels van de spelling van toepassing zijn, spellen ze even goed als de GSZL's, maar tegen de onregelmatige woorden maken ze meer fouten door die woorden te behandelen alsof ze regelmatig zijn; ze schrijven immers vaker fonetisch. Terwijl de GSZL heeft opgemerkt dat THEATER een H bevat, maar misschien niet meer weet waar en daarom THEATHER schrijft of TEATHER, zal dat de ZSGL zijn ontgaan en hij zal het woord als TEATER of TEJATER schrijven.

Over de indeling in groepen moet overigens worden opgemerkt dat zij niet bedoeld is om een typologie met discrete klassen van leerlingen te ontwerpen. Werkwijze en bedoelingen komen overeen met de gangbare praktijk om personen aan het ene einde van een vaardigheidscontinuüm te vergelijken met die aan het andere einde, bijvoorbeeld goede lezers en zwakke lezers, met het doel na te gaan of de zo gevormde categorieën verschillen op andere relevante variabelen. Het verschil is slechts dat er hier twee continua in het geding zijn in plaats van een. De uitkomsten laten zien dat de onderscheiden groepen inderdaad verschillen in hun gemiddelde waarden op andere variabelen. Vooralsnog zijn er echter geen aanwijzingen dat dat te wijten is aan psychologische mechanismen die van elkaar verschillen, integendeel, de (gepoolde binnengroepen) correlatie tusen de nauwkeurigheid bij het spellen van onregelmatige woorden en het lezen van pseudowoorden, 0.46, doet vermoeden dat het mechanisme dat ten grondslag ligt aan het onderscheid tussen de groepen ook binnen elke groep werkzaam is.

De redenering die we volgden is, hoewel in overeenstemming met de feiten, op een aantal punten speculatief. Nadere toetsing is nodig, bijvoorbeeld van de veronderstelling dat de betere leesprestaties van de ZSGL's te danken zijn aan hun gebruik van contextuele informatie. Gaan we er echter vanuit dat de redenering juist is, dan laten zich enkele consequenties van didactische aard formuleren die zich lenen voor toetsing van de hier beschreven gedachtengang.

Op de eerste plaats is er een manier van lezen, "by partial cues", die adequaat kan lijken en dat voor veel leessituaties wellicht ook is, maar die minder rendement geeft waar het de verwerving van spellinginformatie betreft. Als de vraag zich voordoet hoe het spellingonderwijs voor de achterblijvers er dient uit te zien, dan zou men kunnen overwegen manieren van lezen te introduceren die het nodig maken alle orthografische bijzonderheden van woorden in zich op te nemen, bijvoorbeeld door het geven van overschrijfoefeningen (Van Doorn-van Eijsden, 1984) of door fouten te laten opsporen in teksten. Overigens hoeft daarbij niet het enige doel verbetering van spellingkennis te zijn. Verbetering van de leesvaardigheid zou ook doel op zich kunnen zijn. Bryant en Bradley (1980) suggeren dat kinderen nogal eens leesproblemen krijgen rond de leeftijd waarop ze het voortgezet onderwijs ingaan omdat ze dan met nieuwe on-

derwerpen, woorden en contexten te maken krijgen. Als ze lezen door woordherkenning op grond van partiële informatie en contextgebruik, dan voldoet dat niet meer (cfr. Freeborne & Byrne, 1988, voor een vergelijkbare argumentatie m.b.t. jongere kinderen). Bevorderen van het lezen met een fonologische strategie zou een remedie kunnen zijn.

Op de tweede plaats kan men zich afvragen hoe men de zwakke leesprestaties van de GSZL's dient op te vatten. Ze lezen op de cloze-toets immers slechter dan hun brugklasgenoten. Hoewel ze in staat zijn tot preciezer woord-identificatie dan de ZSGL's, lijken ze daarvan minder gebruik te maken om vast te stellen wat er op een bepaalde plaats in een zin kan of moet staan. Een mogelijkheid is dat zij de woorden in de rest van de zin weliswaar nauwkeuriger identificeren, maar ten koste van zoveel verwerkingscapaciteit dat interpretatie-processen van hogere orde niet meer goed mogelijk zijn. De manier waarop GSZL's lezen zou dan weliswaar een manier zijn om een betere speller te worden, maar teveel verwerkingscapaciteit vragen en adequaat tekstbegrip belemmeren. Dat zou het geval kunnen zijn als de fonologische strategie *op zich* het interpreteren van een tekst belemmert, bijvoorbeeld door het bezetten van de repetitie-capaciteit of van andere STM-faciliteiten, of *door de specifieke manier* waarop de GSZL's de fonologische werkwijze uitvoeren. Die is misschien te weinig geautomatiseerd en daardoor te belastend. Als de fonologische strategie op zich tekstbegrip belemmert dan zullen de didactische consequenties afhankelijk zijn van het relatieve gewicht dat men hecht aan tekstbegrip en spellingvaardigheid. Als er echter sprake is van een in principe doeltreffende manier van lezen die te weinig geautomatiseerd is, dan zou men de voorkeur dienen te geven aan leesoefeningen die automatisering beogen, bijvoorbeeld 'flash cards' (Van Bon, Van Kessel & Kortenhorst, 1987).

Het onderzoek maakt duidelijk dat er onder de brugklassers leerlingen te onderscheiden zijn waarvan de lees- en spellingscores niet met elkaar overeenkomen. Sommigen van hen lezen beter dan ze spellen, anderen spellen beter dan ze lezen. Er zijn sterke aanwijzingen dat ze ook van elkaar verschillen in de manier waarop ze lezen en spellen. De uitkomsten suggereren wat de aard van het verschil zou kunnen zijn, maar zoals gebruikelijk werpt het onderzoek een aantal nieuwe vragen op. Beantwoording van die vragen is van belang om de didactische beslissingen die ermee verbonden zijn en dat des te meer omdat er een verband is met het probleem van de dyslectische verschijnselen in het VO: Onder wie vinden we de dyslectici? Onder de ZSGL's, de GSZL's, onder beide of onder de ZSZL's?

Noten

1. Dit is geen gevolg van toetsonbetrouwbaarheid. De KR20 van de cloze-toets was .81, die van de spellingtoets .85; de split-half-coëfficiënten waren zelfs hoger. Ook na correctie voor onbetrouwbaarheid is de samenhang nog verre van volledig: .65.
2. Bij deze toets, ontworpen als groepsgewijs af te nemen alternatief voor de Een-Minuut-Test, moeten zoveel mogelijk woorden worden gelezen in een minuut, met de opdracht de woorden door te strepen die tot een bepaalde categorie (dierennamen) behoren.

Literatuur

Aarle, E.J.M. van & M. Vollebergh, Raders en spellers: Wat is de betekenis van dit onderscheid in de groep zwakke lezers. *Pedagogische Studiën*, 1986, *63*, 339-346.

Aarnoutse, C.A.J., M.J.C. Mommers, B.W.G.M. Smits & J.F.J. van Leeuwe, De ontwikkeling en samenhang van technisch lezen, begrijpend lezen en spellen. *Pedagogische Studiën*, 1986, *63*, 97-110.

Boland, Th. & M.J.C. Mommers, *Development of reading skills in Dutch primary schools: outcomes and prospects*. Paper voor AERA-congres, San Francisco, 1986.

Bon, W.H.J. van, *Aspecten van taalvaardigheid*. Lisse: Swets & Zeitlinger, 1984.

Bon, W.H.J. van, A.E.G. van Kessel & E.P.M. Kortenhorst, Beïnvloeding van woordherkenningssnelheid door middel van flash cards. In: J. Hamers & A. van der Leij (Red.), *Dyslexie 87*. Lisse: Swets & Zeitlinger, 1986.

Bormuth, J.R., Factor validity of cloze tests as measures of reading comprehension. *Reading Research Quarterly*, 1969, *4*, 358-365.

Bouwens, F.G.M., *Strategiegebruik en verschillen in leesvaardigheid en schrijfvaardigheid*. Nijmegen: Doctoraalscriptie K.U.N., 1988.

Bryant, P.E. & L. Bradley, Why some children sometimes write words they don't read. In: U. Frith (Ed), *Cognitive processes in spelling*. London: Academic Press, 1980.

Doorn-van Eijsden, M. van, Leer je spellen door veel te lezen? *Tijdschrift voor Taalbeheersing*, 1984, *6*, 252-263.

Ellis, A.W., *Reading, writing and dyslexia: a cognitive analysis*. London: Lawrence Erlbaum, 1984.

Freeborne, P. & B. Byrne, Wordreading strategies in elementary school children: Relations to comprehension, reading time, and phonemic awarenes. *Reading Research Quarterly*, 1988, *23*, 441-453.

Frith, U., Unexpected spelling problems. In: U. Frith (Ed), *Cognitive processes in spelling*. London: Academic Press, 1980.

Frith, U., Beneath the surface of developmental dyslexia. In: K. Patterson, M. Coltheart & J. Marshall (Eds.), *Surface Dyslexia*. London: Lawrence Erlbaum, 1985.

Geerling, M. & H. Geurts, *Van aap tot hiërarchie: de ontwikkeling van een spellingtoets voor Basis- en Voortgezet onderwijs*. Nijmegen: Doctoraalscriptie K.U.N., 1986.

Henk, W.A., A response to Shanahan, Kamil and Tobin: The case is not yet clozed. *Reading Research Quarterly*, 1982, *17*, 591-595.

Malmquist, E., *Factors related to reading disabilities in het first grade of the elementary school*. Stockholm: Malmquist & Wiksell, 1958.

Rost, D.H., Reading comprehension: skill or skills? *Journal of Research in Reading*, 1989, *12*, 87-113.

Shanahan, T., Nature of the reading-writing relation: an exploratory multivariate analysis. *Jounal of Educational Psychology*, 1984, *76*, 347-368.

Shanahan, T., M.L. Kamil & A.W. Tobin, Cloze as a measure of intersentential comprehension. *Reading Research Quarterly*, 1982, *17*, 229-255.

Smith, N. & A. Zinc, A cloze-based investigation of reading comprehension as a composite of subskills. *Journal of Reading Behavior*, 1977, *9*, 395-398.

Treiman, R., Individual differences among children in spelling and reading styles. *Journal of Experimental Child Psychology*, 1984, *37*, 463-477.

Curricula vitae

W. H. J. van Bon: zie *Pedagogische Studiën*, 1990, *67*, p. 69.

F. G. M. Bouwens studeerde Nederlands en Duits aan de Lerarenopleiding en Nederlandse Taal- en Letterkunde aan de Katholieke Universiteit van Nijmegen. Naast onderzoek naar lees- en spelprestaties verrichtte hij onderzoek op het gebied van het vreemde-talenonderwijs (Instituut voor Toegepaste Sociale Wetenschappen). Momenteel is hij als onderzoeksmedewerker verbonden aan de Vakgroep Toegepaste Taalkunde en Methodologie van de K.U.N.

Adres: Instituut voor Orthopedagogiek, Katholieke Universiteit Nijmegen, Postbus 9103, 6500 HD Nijmegen

Manuscript aanvaard 12-5-'90

Summary

Bon, W.H.J. van & F.G.M. Bouwens. 'Discordant reading and spelling performance of pupils at lower secondary school.' *Pedagogische Studiën*, 1990, *67*, 317-325.

It has been well established that measured correlations between reading and spelling never reach the maximum. Therefore, in individual cases reading and spelling performance can differ: some people are better at reading than spelling, some other people are better at spelling than reading. The present study investigated whether these differences in reading and spelling performance are based on differences in the use of the phonological and the visual-orthographic strategy in both reading and spelling. Good spellers who are poor readers (mean age 12;11) and poor spellers who are good readers (mean age 13;0) read aloud and wrote regular, irregular and pseudo words. Accuracy in spelling and reading as well and reaction time in reading were measured. As indicated on reading aloud pseudo words, good spellers who are poor readers seem to be more skilled in using the phonological strategy. This attention to full cues during reading provides detailed orthographic information which makes them more successful in spelling irregular words. The level of reading performance of poor spellers who are good readers, as measured in a cloze procedure, might indicate a more successful use of contextual information. This strategy however, provides less orthographic information.

PEDON-Conferentie 'Het oog op agogisch onderzoek'

Op 2 februari vond aan de Rijksuniversiteit van Utrecht het eerste congres plaats van PEDON, de Stichting Pedagogisch Onderzoek i.o. die onderdeel vormt van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Na een lange voorgeschiedenis waarin verschillende pogingen zijn ondernomen om een stichting van de grond te krijgen, is het in 1988 gelukt om voor een proeftijd van een paar jaar, de status van een stichting 'in oprichting' Bij NWO te verkrijgen. Gegeven het enthousiasme waarmee sindsdien gewerkt is aan onderzoeksvoorstellen en symposia in de diverse werkgemeenschappen en nu dan een congres van de stichting als geheel, en de reactie daarop van NWO-zijde, ziet het er naar uit dat PEDON een volwaardig lid zal worden van het stichtingenbestand dat NWO kent. Onderzoekers kunnen hun voorstellen indienen bij een van de diverse werkgemeenschappen: wijsgerige en historische pedagogiek, orthopedagogiek, andragologisch en sociaal-pedagogisch/jeugdonderzoek, onderwijsonderzoek, en sinds kort ook empirisch gezinspedagogisch onderzoek.

Het congres droeg na rijp beraad de titel 'Het oog op agogisch onderzoek', waaruit blijkt hoeveel rekening er wordt gehouden met de andragologische minderheid in de overwegend pedagogisch gekleurde bijt. De bedoeling van het congres was om een aantal zaken die voor PEDON als geheel van belang zijn te presenteren en daarnaast de verschillende werkgemeenschappen de kans te geven een drietal sprekers aan het woord te laten. Bovendien hadden de, veelal jongere, onderzoekers de gelegenheid om een posterpresentatie te verzorgen. Ten slotte kwamen medewerkers van NWO aan het woord teneinde de verschillende subsidiemogelijkheden te bespreken. Het congres was derhalve zeer breed van opzet, een doorsnee van hetgeen in PEDON aan de orde is op dit moment. Een dergelijke opzet brengt het gevaar met zich mee dat de onderwerpen te gevarieerd zijn voor bezoekers die slechts geïnteresseerd zijn in een bepaald onderzoeksdomein, aangezien zelfs in de parallelsessies verzorgd door de werkgemeenschappen, de onderwerpen sterk uiteenliepen. Anderzijds werd de bezoekers wel de kans geboden om kennis te nemen van hetgeen er in andere deeldisciplines gebeurt, wat er tegenwoordig door de typische produktiedruk en de inhoudelijke beperkingen die voorwaardelijk gefinancierde onderzoeksprogramma's kenmerken, gewoonlijk niet van komt.

In de ochtend vond een tweetal lezingen plaats. Creemers (Instituut voor Onderwijsonderzoek, RION, Groningen) gaf een exposé over de ontwikkeling binnen de pedagogische deeldisciplines in de afgelopen tien jaar onder de titel 'Ouder en wijzer?' Hij concludeerde o.a., dat het aantal deeldisciplines en leerstoelen is gereduceerd en dat er veel energie gestoken is in bestuurlijke problemen (op facultair en vakgroepsniveau) met als gevolg negatieve effecten voor de wetenschapsbeoefening. Daar tegenover staat dat de ontwikkeling van de organisatie binnen NWO op gang is gekomen – PEDON – en dat de produktie van wetenschappelijke publikaties binnen het verband van de voorwaardelijk gefinancierde programma's groter is geworden. 'De discipline bestaat nog', aldus Creemers, er is meer realisme en minder wetenschapstheoretische discussie dan tien jaar geleden. Vervolgens besprak hij de ontwikkelingen in de deeldisciplines die in deze kroniek zullen worden overgeslagen omdat die anders te lang zou worden. Ten slotte kwam hij tot een aantal aanbevelingen, zoals: meer interdisciplinaire samenwerking, minder onderwerpen voor theorievorming en onderzoek, en minder directe praktijk- en beleidsgerichtheid (wellicht kan hij dit laatste ook eens wijs maken aan het Instituut voor Onderzoek van het Onderwijs, SVO, dat in haar programma's juist beleid en praktijk predikt en wat er aan theorievorming kan geschieden, ook nog in eigen hand houdt). Daarnaast bepleitte hij dat Wijsgerige en Historische Pedagogiek meer 'moederdisciplinegebonden' zou worden, d.w.z. meer

Pedagogische Studiën 1990 (67) 326-329

georiënteerd op de filosofie, en dat er meer empirische en theoretische gerichtheid binnen (gebieden van) de orthopedagogiek en de onderwijskunde zou komen. Als referent voor het verhaal van Creemers trad Imelman (Rijksuniversiteit, Utrecht) op die concludeerde dat Creemers sterk bevooroordeeld was met een voorkeur voor een bepaald type onderzoek. Tevens, dat dit soort onderzoek reeds verricht werd in een aantal andere NWO-stichtingen, en dat, wanneer Creemers' ideeën serieus werden genomen, het pedagogisch onderzoek wel opgeheven kon worden. Dit, voor de kenners van Imelmans werkwijze niet, maar voor anderen wellicht wel verrassende standpunt, bleek niet de juiste manier om een discussie met de zaal te stimuleren. Duidelijk is echter wel dat de opmerkingen die er gemaakt werden m.b.t. de diverse deeldisciplines, in de wandelgang heel wat stof tot discussie gaven, zodat geconcludeerd kan worden dat de ondankbare taak die Creemers op zich genomen heeft om een overzicht te geven van ontwikkelingen in een wetenschapsterrein dat zo divers is als de pedagogiek, toch vrucht heeft afgeworpen.

Vervolgens was het woord aan Reitsma (Paedologisch Instituut, Amsterdam) die samen met een aantal anderen, een voorstel voor een aandachtsgebied binnen NWO had ingediend via PEDON onder de titel 'Functional literacy'. Een aandachtsgebied is, in tegenstelling tot de gebruikelijke voorstellen die één project betreffen, een onderzoeksprogramma met voorstellen voor samenhangende projectvoorstellen. Reitsma begon zijn toelichting in het Engels, daarmee niet alleen demonstrerend dat internationalisering hem ter harte gaat, maar vooral ook dat NWO-procedures niet altijd even helder en consequent zijn. Bij het schrijven van het voorstel was hem immers te verstaan gegeven dat het in het Engels moest omdat het voorgelegd zou worden aan internationale deskundigen op het terrein. Echter, daar was niets van terecht gekomen omdat de besluitvorming geheel binnen de Nederlandse huiskamers kon plaatsvinden. Na eerst 'ja' op het niveau van PEDON, kregen de indieners 'nee' op het niveau van het overkoepelende gebiedsbestuur voor de maatschappij- en gedragswetenschappen en vervolgens weer 'ja' op het niveau van het NWO-bestuur.

Het aandachtsgebied m.b.t. functionele geletterdheid beoogt onderzoek op drie terreinen: 1) beschrijving en diagnostiek van functionele geletterdheid, 2) factoren die ten grondslag liggen aan of mede bijdragen aan de ontwikkeling van (on-)geletterdheid en 3) effectieve interventiemethoden.

Reitsma gaf vervolgens een uitwerking van de eerste twee terreinen. Ten eerste zou een hoeveelheid toetsen voor functionele geletterdheid ontwikkeld en beproefd moeten worden die te maken hebben met de vaardigheid zoals die in het dagelijkse leven vereist wordt. Op basis van rationele analyse kan een taxonomie van taken ontstaan die ten grondslag ligt aan de ontwikkeling van toetsen. Afname van deze toetsen op diverse leefdtijdsniveaus kan een indruk geven van de mate van functionele (on-)geletterdheid. Ten tweede, de ontwikkeling van de leesvaardigheid in het basisonderwijs zou onderzocht moeten worden, waarbij alle mogelijke aspecten erbij betrokken worden: leerling-, leerkracht-, leerinhoud- en instructievariabelen, klasseorganisatie en de interacties daartussen. Het plan, waarvan de verkorte vorm toegestuurd was aan alle congresdeelnemers, werd becommentarieerd door Mommers (Katholieke Universiteit, Nijmegen) en De Glopper (Stichting Centrum voor Onderzoek van het Onderwijs, SCO, Amsterdam). Eerstgenoemde sprak zijn waardering voor het voorstel uit maar miste een uitwerking naar het schrijfgedrag. Voorts kwam hij tot de aanbeveling om, ook al gaat het om vaardigheden die maatschappelijk zeer relevant zijn, toch de wetenschappelijke weg te volgen en de nadruk sterk op theorievorming te leggen. Ten slotte gaf hij een voorbeeld van een taxonomie van de vaardigheden die een lezer tot een geoefende lezer maken en bepleitte om in het kiezen van meetinstrumenten gebruik te maken van de diverse mogelijkheden die in uiteenlopende instituten voor handen zijn. De Glopper wees vervolgens in dit verband op de internationale vergelijkende studie die er naar functionele geletterdheid ondernomen wordt in Europa en naar het werk in eigen land, o.a. als onderdeel en gevolg van de proefstudie peilingsonderzoek. Op grond van de ervaringen bepleitte hij m.b.t. de functionele leestaken die onderzocht gaan worden, een normatief standpunt (welke taken achten wij van belang om te beheersen?), en, in het verlengde daarvan, van een theoretische analyse van de taakkenmer-

ken en -eisen zodat de taxonomie gemaakt kan worden. Voor het onderzoek naar de ontwikkeling van de leesvaardigheid zag hij als bedreiging dat het teveel zou worden en daardoor, qua resultaat, te weinig, omdat leesvaardigheid slaat op een grote diversiteit aan inhouden, bijv. narratieve versus zakelijke teksten, fictie of non-fictie, zinnen of tabellen, etc.

Na de lunch verspreidden de deelnemers zich over de parallelsessies die waren georganiseerd door de werkgemeenschappen. In deze kroniek zal volstaan worden met het noemen van de thema's die aan de orde zijn gesteld.

De Werkgemeenschap Andragologisch en Sociaal-Pedagogisch/Jeugdonderzoek presenteerde zich middels drie voordrachten. Polstra (Rijksuniversiteit Groningen) vroeg zich af: 'Is een evaluator zowel een onderzoeker als een ontwikkelaar?', toegespitst op de arbeidstherapie in de geestelijke gezondheidszorg. Du Bois-Reymond, Hazekamp en Meeus (resp. Rijksuniversiteit Leiden, Vrije Universiteit Amsterdam en Rijksuniversiteit Utrecht) bespraken een drietal 'Veranderende ontwikkelingstaken in de adolescentie': het met goed gevolg doorlopen van de schoolloopbaan, het aangaan van intieme relaties en het vinden van een levensovertuiging of ideologie. Van Haaster en Snel (Universiteit van Amsterdam) hadden het 'Over gebruikerstaal' die geconstrueerd wordt teneinde psychotische patiënten in staat te stellen beter te communiceren met hun omgeving.

De Werkgemeenschap Onderwijsonderzoek had in haar sessie Tillema (Rijksuniversiteit Groningen) over 'Training en coaching van bekwaamheden van leerkrachten', gericht op conceptueel leren in relatie tot praktische uitoefening van vaardigheden en transfer naar de klas. Streumer en Doornekamp (Universiteit Twente) gaven een exposé over 'Het ontwerpen van minicursussen voor techniek', waarbij twee vormen werden gecontrasteerd: demonstratie van een praktische vaardigheid middels een geschreven tekst of audio-visueel. Roeders en Van Oudenhoven (Rijksuniversiteit Groningen) hielden een voordracht over 'Schoolklasorganisatie en effectiviteit in samenwerkend leren', vooral m.b.t. de condities waaronder samenwerkend leren effectief kan plaatsvinden in het onderwijs.

De Werkgemeenschap Orthopedagogiek, presenteerde bij monde van Hellendoorn (Rijksuniversiteit Leiden) het 'Onderzoek naar effecten van beeldcommunicatie', waarin vooral de methodologische overwegingen bij het op te starten evaluatie-onderzoek naar deze therapie bij kinderen werden toegelicht. Roelofs (Katholieke Universiteit Nijmegen) besprak in verband met 'De ernst van lichamelijke kindermishandeling' de factoren die bepalend zijn voor de ernst van de fysieke mishandeling. Van der Leij (Vrije Universiteit Amsterdam) lichtte de voorstudie toe van het project 'Directe woordherkenning bij kinderen met leesproblemen': in hoeverre zijn leerlingen met ernstige leesproblemen tot directe woordherkenning in staat?

De Werkgemeenschap Wijsgerige en Historische Pedagogiek was vertegenwoordigd door Tellings (Katholieke Universiteit Nijmegen) die het had over 'Doven-taalontwikkeling: een onderzoek naar grondslagen van de dovenopvoeding', toegespitst op de strijd tussen de 'oralisten' en de 'totalisten' die zich onderscheiden in de mate waarin naast orale taal ook manuele communicatie is toegestaan. Kuipers (Universiteit van Amsterdam) schetste 'Een theoretisch kader voor onderzoek naar de ontstaansgeschiedenis van de Werkplaats Kindergemeenschap' van Kees en Betty Boeke. Biesta (Rijksuniversiteit Leiden) hield een voordracht over 'Transactie, ervaring en kennen: het cultureel naturalisme van John Dewey', met name Dewey's visie op kennen als modus van de ervaring die betrekking heeft op de voortgang van het handelen.

Vervolgens kregen onderzoekers de kans om middels een posterpresentatie ruchtbaarheid aan hun onderzoek te geven. Van de 27 posterpresentaties kwamen er 12 uit de orthopedagogische hoek, waarvan er 10 betrekking hadden op onderwijsleerproblemen/leerstoornissen. Andragologie en jeugdstudies waren vertegenwoordigd met 5 onderzoeken, gezinspedagogiek ook, wijsgerig/historisch onderzoek met 2 en onderwijsonderzoek met 3. De onderwerpen varieerden sterk en gaven, tesamen met de voordrachten in de sessies, een aardig beeld van de breedheid van de problematiek waarmee pedagogische en andragogische onderzoekers zich bezighouden. Voor degenen die een overzicht willen hebben van de onderwerpen, er is een congresboek in voorbereiding dat over een aantal maanden bij ACCO zal verschijnen. Het zal niet alleen

de neerslag van de toespraken en voordrachten bevatten, maar ook van de postersessies.

De dag werd besloten door een aantal NWO-medewerkers die de verschillende vormen van subsidie toelichtten. Geïnteresseerden kunnen zich voor algemene informatie het beste wenden tot de Nederlandse Organisatie voor Wetenschappelijk Onderzoek. Voor informatie over de Stichting Pedagogisch Onderzoek kunt u op hetzelfde adres terecht: Postbus 93138, 2509 AC 's-Gravenhage.

Al met al een congres waarmee PEDON heeft laten zien dat het levensvatbaar is – er waren ruim 150 belangstellenden – en in staat om ook in deze vorm een forum voor presentatie en uitwisseling van wetenschappelijke kennis te organiseren. Tevens dat er een boeiend scala aan onderzoek onder de PEDON-vlag kan worden gerangschikt. Binnen de stichting en de werkgemeenschappen zal in de naaste toekomst moeten worden aangetoond dat blijvend aan de kwaliteitseisen voor wetenschappelijk onderzoek kan worden voldaan. Voor subsidiëring van dergelijk onderzoek door NWO is tot nu toe een bescheiden budget beschikbaar (in 1989 werd al met al zo'n 750.000 gulden aan pedagogisch onderzoek besteed). Hoopgevend is echter dat in enige groei is voorzien.

A. van der Leij

Boekbespreking

H. van den Bergh, *Examens geëxamineerd: Een onderzoek naar de validiteit van examens Nederlands voor LBO en MAVO.* Het Instituut voor Onderzoek van het Onderwijs S.V.O., 's-Gravenhage, 1988, 198 pag., ISBN 90 6472 139 4.

Dit boek bevat een onderzoeksverslag van werkzaamheden in het kader van een door SVO gesubsidieerd project 'Examenproblematiek in het Voortgezet Onderwijs'. Het werd begin 1989 als proefschrift verdedigd aan de Universiteit van Amsterdam. Het doel van dit project was een evaluatie van de doorgevoerde veranderingen in het Centraal Schriftelijk Examen (CSE) Nederlands voor LBO- en MAVO-kandidaten. In dit proefschrift wordt verslag gedaan van het deelproject betreffende de validiteit van de nieuwe examens. Deze nieuwe Centraal Schriftelijke Examens Nederlands voor LBO- en MAVO-leerlingen kennen twee belangrijke wijzigingen: 1) voor het onderdeel tekstbegrip wordt het cijfer van de kandidaat voor 70% bepaald door antwoorden op gesloten vragen en voor 30% door die op open vragen; 2) in het onderdeel schrijfvaardigheden zijn drie à vier functionele opdrachten opgenomen, schrijfopdrachten gebaseerd op taalgebruiks-situaties. Deze veranderingen in het CSE vormden de aanleiding tot twee vragen die in het onderzoek centraal staan: 1. doet de beantwoording van open en gesloten vragen tekstbegrip een beroep op dezelfde intellectuele vaardigheid, en 2. wordt bij het maken van functionele schrijfopdrachten en bij het schrijven van een opstel een beroep gedaan op dezelfde intellectuele vaardigheden?

De opbouw van het boek is als volgt. Na de inleiding in het eerste hoofdstuk wordt in hoofdstuk 2 verslag gedaan van de constructie van het instrumentarium nl. vier tekstbegriptoetsen, de schrijfvaardigheidsinstrumenten en de intellectuele vaardigheidstoetsen. Het meten van deze intellectuele vaardigheden is immers nodig teneinde na te gaan of op open versus gesloten vragen tekstbegrip een beroep doen op dezelfde dan wel verschillende intellectuele vaardigheden. Hetzelfde geldt voor het onderscheid tussen het maken van een functionele schrijfopdracht en het schrijven van een traditioneel opstel. Uitgangspunt voor deze intellectuele vaardigheidstoetsen wordt gevormd door Guilfords Structure-of-Intellect model. Dit SI-model gaat van de veronderstelling uit dat elke vaardigheid gekenmerkt kan worden op drie dimensies n.l. operaties, inhouden en produkten. Het model kent 5 operaties (b.v. evaluatie, convergente produktie), 5 inhouden (b.v. gedragsmatig, semantisch) en 6 organisaties (b.v. relaties, transformaties). Hiermee wordt een taxonomie van vaardigheden opgegeven. Probleemtypen die verschillen op een van de drie dimensies doen een beroep op verschillende disposities. Prestaties op taken die niet verschillen op één van de drie dimensies hangen nauwer met elkaar samen, dan die wel verschillen op één van de drie dimensies. Niet alle mogelijke $5 \times 5 \times 6 = 150$ SI-vaardigheden bleken in het onderzoek van belang. Op grond van literatuurstudie, een analyse van tekstbegripvragen en een analyse van hardop-schrijf protocollen beide aan de hand van de semantische vaardigheden uit het SI-model, werd het aantal relevante intellectuele vaardigheden drastisch ingeperkt. Uiteindelijk viel de keuze op 16 semantische SI-vaardigheden voor elk waarvan drie toetsen geconstrueerd werden. Dit aantal toetsen ($3 \times 16 = 48$) was nog te groot en er werd dan ook naar een verdere beperking gezocht van het aantal items. Het laatste kwam tot stand door de constructie van zgn. Rasch-homogene vaardigheidstoetsen.

De hoofdstukken 3, 5 en 6 geven de resultaten van de analyses van de leerlingprestaties op de leesvaardigheidstoetsen, de vijf schrijfvaardigheidstoetsen en de 47 geconstrueerde intellectuele vaardigheidstoetsen. De relatie tussen de SI-vaardigheden en de leesvaardigheid gemeten met open en gesloten vragen komt in hoofdstuk 7 aan de orde, die tussen de SI-vaardigheden en de schrijfvaardigheden in hoofdstuk 8.

Een aparte plaats wordt ingeruimd voor het probleem van de psychometrische betrouwbaarheid van beoordelaars. Dit is het onderwerp van hoofdstuk 4. Het is vooral van belang wanneer produkten worden beoordeeld als b.v. opstellen. Drie modellen worden behandeld: een model voor congenerieke oordelen, voor tau-equivalente oordelen en voor parallelle oordelen. Hoofdstuk 9 bevat enkele exploraties, de behandeling van een aantal methodologische problemen en als afsluiting enige onderwijskundige conclusies die op basis van het onderzoek van belang zijn voor de examenpraktijk.

Strikt genomen geven hoofdstuk 7 en 8 antwoord op beide onderzoeksvragen. Uit hoofdstuk 7 blijkt een aanzienlijk deel van de variantie in de leesvaardigheidsprestaties verklaard te kunnen worden uit de gemeten SI-vaardigheden. De toetsen tekstbegrip met open en gesloten vragen zijn congeneriek met betrekking tot de gemeten SI-vaardigheden; de beantwoording van open en gesloten vragen tekst-

Pedagogische Studiën 1990 (67) 330-339

begrip doet een beroep op dezelfde intellectuele vaardigheden. De belangrijkste conclusie die uit hoofdstuk 8 is te trekken is, dat bij het maken van een functionele schrijfopdracht en het schrijven van een traditioneel opstel dezelfde intellectuele vaardigheden zijn betrokken, vooropgesteld dat de schrijfprodukten op dezelfde wijze beoordeeld worden. Het is duidelijk dat de beantwoording van beide onderzoeksvragen eerst mogelijk was na de constructie van het instrumentarium, de analyse van de leesvaardigheidstoetsen, de schrijfvaardigheidstoetsen en de intellectuele vaardigheidstoetsen. En lijkt de beantwoording van beide onderzoeksvragen eenvoudig te geven te zijn, niets is minder waar. Om tot het antwoord op de eerste vraag te komen is heel zorgvuldig te werk gegaan. Er is eerst onderzocht of er een verschil is aan te tonen in gemiddelde prestaties en varianties op de tekstbegriptoets die toe te schrijven is aan de vraagvorm. Dit bleek het geval te zijn, en voor deze verschillen is een verklaring gezocht. Nu zeggen verschillen in (waargenomen) gemiddelden en varianties weinig over verschillen in intellectuele vaardigheden die nodig zijn voor de vragen tekstbegrip. Hiervoor was nodig een nadere analyse van de correlaties tussen de prestaties van de leerlingen op de toetsen (met open en gesloten vragen) tekstbegrip en hun prestaties op de intellectuele vaardigheidstoetsen. Op grond van theoretische overwegingen werden negen psychometrische modellen geformuleerd met betrekking tot de structuur van intellectuele vaardigheden. Eén ervan is het SI-model. In dit onderzoek worden zestien vaardigheden onderscheiden, vijftien worden elk geïndiceerd door drie toetsen en één vaardigheid door twee. De correlaties tussen de scores op de 47 intellectuele vaardigheidstoetsen zouden theoretisch een structuur moeten vertonen conform het SI-model. Dit werd onderzocht en daarnaast ook in hoeverre deze structuur wellicht conform andere modellen – acht in getal – van intellectuele vaardigheden is. Uit de analyse bleken modellen met een structuur overeenkomstig het SI-model het best te passen bij de structuur van de geobserveerdercorrelaties. Door bestudering, ten slotte, van de regressies van de begrijpend leesvaardigheidscores op de intellectuele vaardigheidsscores kwam het antwoord op de eerste onderzoeksvraag. Een analoge werkwijze werd gevolgd voor de beantwoording van de tweede onderzoeksvraag.

Het geheel overziend kan geconcludeerd worden dat het onderzoek zeer zorgvuldig is opgezet en uitgewerkt. Er is aandacht besteed aan de ontwikkeling van het instrumentarium, aan de theoretische inkadering van zowel tekstbegrip en schrijfvaardigheid als intellectuele vaardigheden, en aan het betrouwbaarheidsprobleem. De data-analyse is nauwgezet uitgevoerd waarbij gebruik is gemaakt van geavanceerde technieken zoals structurele vergelijkingsmodellen, Rasch-analyse, logit-analyse. Het SI-model van Guilford is van stal gehaald om de gebruikte toetsen tegen te valideren. Je zou ook, andersom, kunnen zeggen dat met dit onderzoek het SI-model is gevalideerd, althans een plakje van het volledige model. Want wie zou nog in staat zijn om in een onderzoek het complete SI-model te valideren. Hoe het ook zij, alles bij elkaar genomen vormt deze dissertatie een degelijk en deugdelijk stuk onderwijskundig onderzoek. De hoofdstukken 2 t/m 8 laten zich lezen als afzonderlijke stukjes. Dit is ook niet verwonderlijk: elk van deze hoofdstukken is als tijdschriftartikel verschenen of aangeboden voor plaatsing. Omdat deze hoofdstukken min of meer op zichzelf staan mist het boek als zodanig de coherente en consequente opbouw van een monografie. Herhalingen b.v. zijn onvermijdelijk, de auteur merkt dit overigens zelf ook op.

Naast alle waardering voor het werk zijn er ook kritische kanttekeningen te maken. In de eerste plaats is daar de keuze van het SI-model met de 150 vaardigheden, Ockhams scheermes: *entia non sunt multiplicanda praeter necessitam* schijnt Guilford geheel vreemd te zijn geweest. Bestond het SI-model aan het eind van de zeventiger jaren nog uit 120 intellectuele vaardigheden, het begin van de tachtiger jaren bracht een uitbreiding naar 150. Spaarzaamheid (parsimonia) is ver te zoeken. Hoe het ook zij, nog lang niet alle 150 vaardigheden zijn op bevredigende empirische wijze aangetoond. Los hiervan is er kritiek op de wijze waarop het SI-model doorgaans gevalideerd wordt n.l. met behulp van een of andere vorm van factor-analyse al dan niet gevolgd door een rotatie. Nu werd in het onderzoek voor de data-analyse gebruik gemaakt van LISREL, niettemin blijft de vraag of andere structurele vergelijkingsmodellen de data niet even goed hadden kunnen beschrijven c.q. verklaren. Dit is de vraag naar de klasse van equivalente modellen voor een gegeven databestand (zie b.v. Luijben, 1989). Een punt van kritiek van in zekere zin andere aard is die vanuit het gezichtspunt van de cognitieve psychologie. In een notedop komt deze kritiek op psychometrische modellen op het volgende neer. Dergelijke modellen lijken geen inhoudelijke psychologische rechtvaardiging te geven voor hetgeen zich op itemniveau afspeelt. Hoe lost een persoon een item van een test op? Dit is een vraag waarop geen antwoord wordt gegeven. Evenmin wordt antwoord gegeven op de vraag hoe de psychologie van de test als geheel eruit ziet. Validering van de betekenis van de test is externe validering buiten het psychometrische model zelf. Voorts wordt de psychometrische modellen verweten – meer hierover is te vinden bij Snow en Lohman (1989) – simplificerende assumpties te maken (b.v. ongecorreleerde fouten, en stochastische onafhankelijkheid c.q. unidimensionaliteit bij item-responsmodellen). Uit het besproken onderzoek wordt niet goed duidelijk welke cognitieve processen ten grondslag liggen aan het maken van de gebruikte SI-tests. Deze kritiek geldt het SI-model als zodanig, en niet zozeer het onderzoek. Dit

stelt zich immers ten doel de validiteit van examens Nederlands voor LBO en MAVO na te gaan. Een andere kritische kanttekening betreft de wijze waarop onderzocht wordt of de toetsen als unidimensioneel zijn te beschouwen. Dit wordt gedaan via principale componentenanalyse op de tetrachorische correlaties tussen de items van een toets. Afgezien van het feit dat een matrix van tetrachorische correlaties negatieve eigenwaarden kan opleveren, is het de vraag of een verklaarde variantie van 23% tot 54% een noodzakelijke en voldoende voorwaarde is voor unidimensionaliteit (zie Hattie, 1985 en de daarin aangehaalde literatuur). Voorts zijn er nog de detailopmerkingen waarvan er een is: waarom de voorkeur gegeven wordt aan logitmodellen boven loglineaire modellen voor de analyse van correcte antwoorden (hoofdstuk 3). Het voordeel van toepassing van laatstgenoemde modellen is dat zij structurele verbanden tussen verklarende variabelen blootleggen. Een loglineaire analyse van deze proporties had ook aangevuld kunnen worden door een correspondentie-analyse. Deze laatste analyse had wellicht aard en structuur van de residuen 'zichtbaarder' gemaakt (vgl. Van der Heijden, De Falguerolles en De Leeuw, 1989 voor een overzicht).

Al met al blijft *Examens geëxamineerd* een goed boek. Dat er niettemin kanttekeningen bij te plaatsen zijn bevestigt alleen maar het feit dat onderzoek nooit af is, en dat op elk empirisch onderzoek wel iets aan te merken is.

L. J. Th. van der Kamp

Literatuur

Hattie, J., Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 1985, *9*, 139-169.

Heijden, P. G. M. van der, A. De Falguerolles, & J. de Leeuw, A combined approach to contingency table analysis using correspondence analysis and log-linear analysis. *Applied Statistics*, 1989, *38*, 249-292.

Luijben, Th. C. W., *Statistical quidance for model modification in covariance structure analysis.* Amsterdam: Sociometric Research Foundation, 1989.

Snow, R. E. & D. F. Lohman Implication of cognitive psychology for educational measurement. In: R. L. Linn (Ed.). *Educational Measurement*, 3rd ed. New York: American Council on Education, 1989.

N. Verloop, *Interactive cognitions of student-teachers. An intervention study.* CITO, Arnhem, 1989, 279 pag., *f* 24,50, ISBN 90-9002900-1.

Op 28 juni 1989 promoveerde N. Verloop aan de Rijksuniversiteit te Leiden cum laude op het onderwerp "Interactive cognitions of student-teachers. An intervention study". In het proefschrift wordt verslag gedaan van een onderzoek naar de effecten van speciaal voor de studie ontwikkelde opleidingsmaterialen op de cognities en het lesgedrag van aanstaande onderwijsgevenden. In concreto gaat het om twee videobanden waarin een bepaalde onderwijskundige theorie aan de hand van concrete lesfragmenten aanschouwelijk is gemaakt. Op de ene band staat het Advance Organizer Model van Ausubel centraal, op de andere band het Role Playing Model van Shaftel.

In het eerste hoofdstuk van de dissertatie wordt de context van de studie toegelicht: de (onbevredigende) relatie tussen theorie en praktijk in het opleidingsonderwijs. In lijn van de cognitieve oriëntatie in het opleidingsonderzoek wordt verondersteld dat wanneer onderwijskundige theorieën niet alleen op verbaal-theoretisch niveau maar ook op het niveau van het observeerbare gedrag worden aangeboden, (elementen van) deze theorieën beter deel gaan uitmaken van de interactieve cognities van aanstaande onderwijsgevenden en tevens hun lesgedrag beïnvloeden.

In hoofdstuk 2 t/m 4 wordt het theoretisch kader van de studie geschetst. Hoofdstuk 2 handelt over de functies die onderwijskundige theorieën kunnen vervullen in het opleidingsonderwijs. Mede op basis van empirische evidentie wordt aannemelijk gemaakt dat deze theorieën, met name wanneer deze worden aangeboden in de vorm van protocolmaterialen, (aanstaande) leraren mogelijkheden bieden het onderwijsleerproces gedifferentieerd waar te nemen en te interpreteren. In hoofdstuk 3 wordt een overzicht gegeven van tot nu toe verricht onderzoek naar cognities van (aanstaande) onderwijsgevenden en van de belangrijkste problemen die zich daarbij kunnen voordoen. Tevens wordt het belang besproken van onderzoek naar de relatie tussen cognities en gedrag. In hoofdstuk 4 wordt uitgebreid ingegaan op de status en de problemen van de stimulated recall methode, welke een belangrijkrijke functie vervult in het onderzoek van Verloop.

In hoofdstuk 5 t/m 7 wordt verslag gedaan van het empirisch onderzoek. In hoofdstuk 5 worden allereerst de onderzoeksvragen gepresenteerd. De hoofdvragen betreffen het effect van het ontwikkelde treatmentmateriaal op de interactieve cognities van aanstaande onderwijsgevenden, op hun lesge-

drag en op de relatie cognitie en gedrag. Omdat vanuit de literatuur verondersteld kan worden dat eventuele experimentele effecten mede bepaald zouden kunnen worden door de waardering die aanstaande leraren hebben voor het onderwijskunde-programma van de opleiding, is hierover eveneens een onderzoeksvraag geformuleerd. Hoofdstuk 6 behandelt de onderzoeksmethode. Uitvoerig wordt ingegaan op de (keuzen gemaakt tijdens de) ontwikkeling van het treatmentmateriaal. Het effect van het materiaal is onderzocht bij 30 Pabo- en NLO-studenten. Tien studenten kregen de onderwijskundige theorie op schrift aangeboden, gevolgd door de videobanden. Tien anderen werden na de schriftelijke instructie op de klassieke, verbaal-theoretische wijze met de theorie geconfronteerd. Een controlegroep van eveneens tien studenten volgde het normale programma van de opleiding. Een maand na de treatment gaven alle studenten twee lessen op hun eigen stageschool, die op videoband werden opgenomen. Na afloop van de lessen werden m.b.v. de stimulated recall-methode de interactieve cognities van de studenten achterhaald. Aan de hand van een speciaal ontworpen categorieënsysteem zijn vervolgens het stimulated recall materiaal en de videoopnamen van de gegeven lessen geanalyseerd, waarbij is nagegaan in welke mate de kernelementen uit de onderwijskundige theorieën terug te vinden waren in resp. de interactieve cognities van de studenten en hun lesgedrag. Het bleek dat belangrijke verschillen aantoonbaar waren tussen de drie groepen aanstaande leraren, in die zin dat de kernelementen uit de theorie significant vaker voorkwamen in de interactieve cognities en in het lesgedrag van de groep die het videomateriaal had doorgewerkt. Ook t.a.v. het verband tussen interactieve cognities en lesgedrag traden, voor zover gemeten op het niveau van de afzonderlijke cognities en gedragingen, belangrijke verschillen op. Bij studenten uit de videogroep ging theorie-consistent gedrag vaker gepaard met interactieve cognities die op de betreffende onderwijskundige theorie betrekking hadden dan bij studenten uit beide andere groepen. De waardering van aanstaande leraren voor onderwijskundige theorie als zodanig bleek als interveniërende variabele geen invloed te hebben. In hoofdstuk 8 ten slotte worden deze resultaten bediscussieerd.

Het geheel overziend kan worden geconcludeerd dat de auteur zonder meer waardering verdient voor de gedegen wijze waarop dit onderzoek is opgezet, uitgevoerd en gerapporteerd. De studie is aanmerkelijk complexer en meer sophisticated dan in een korte samenvatting als hierboven tot uitdrukking kan worden gebracht. Op een heldere en inzichtelijke wijze worden gemaakte keuzen en genomen beslissingen (theoretisch) onderbouwd. Bij de ontwikkeling van het treatmentmateriaal en de diverse onderzoeksinstrumenten is grondige aandacht besteed aan kwaliteitseisen als betrouwbaarheid en validiteit. De dataverwerking heeft zeer zorgvuldig en nauwkeurig plaatsgevonden en de verzamelde onderzoeksdata zijn optimaal gebruikt.

Uiteraard blijven er ook t.a.v. deze studie opmerkingen te maken en vragen te stellen. Juist omdat in het proefschrift zoveel keuzen en beslissingen uitvoerig worden verantwoord, valt bijvoorbeeld op dat geen verantwoording wordt gegeven van de mediakundige principes die zijn gehanteerd bij de ontwikkeling van het videomateriaal. Ook is het jammer – maar gezien de toch al grote omvang van het onderzoeksinstrumentarium goed begrijpelijk – dat mogelijke interacties tussen een bepaalde treatment en bepaalde studentkenmerken niet nader zijn onderzocht. Een vraag die naar aanleiding van het onderzoek rijst is welke aspecten van het treatmentmateriaal nu precies de beoogde effecten tot stand hebben gebracht. Is het louter een kwestie van het zichbaar maken van bepaalde theorie-elementen in concrete gedragingen of worden op die manier andere, voor het handelen van aanstaande onderwijsgevenden meer relevante soorten kennis overgedragen en leerprocessen tot stand gebracht dan via de verbaal-theoretische methode mogelijk is? Een logische volgende stap zou zijn nader te onderzoeken waarom exact het materiaal werkt zoals het werkt.

Dit soort opmerkingen en vragen laat echter onverlet de waarde en het belang van het onderzoek als geheel. Er is door de auteur een grote mate van vakmanschap gedemonstreerd en de theoretische en praktische waarde van deze interventiestudie is groot. Er ligt opleidingsmateriaal en er ligt theorie van waaruit voorspeld kan worden dat het materiaal bepaalde leereffecten tot stand zal brengen.

A.H. Corporaal

Ernst Mulder, *Beginsel en Beroep. Pedagogiek aan de Universiteit in Nederland 1900-1940* (diss.). Amsterdamse Historische Reeks nr 15, Amsterdam, Historische Seminarium van de Universiteit van Amsterdam, 1989, 360 pag., *f* 45,–

Dat een met het Réveil verbonden classicus uit het Nederlandse patriciaat, een natuurkundige uit de geassimileerde joodse bovenlaag in Amsterdam, een afvallige onderwijzer uit een arm calvinistisch éénoudergezin op het Friese platteland, een katholieke priester en filosoof uit een rijke boerenfamilie in de Betuwe en een predikant uit het gezin van een afgescheiden godsdienstonderwijzer in het noorden van Overijssel een gezelschap vormen dat, dwars door scheidingen van klasse en zuil heen snel gelijk gestemd geacht kan worden ligt niet voor de hand.

Toch is dit de slotsom van de buitengemeen boeiende geschiedenis van de "Pedagogiek aan de universiteit in Nederland van 1900 tot 1940" die door Ernst Mulder in een kloek proefschrift werd vastgelegd. Gunning, Kohnstamm, Casimir, Hoogveld en Waterink, ook wel de pedagogen van de eerste generatie genoemd, konden het zowel vakmatig als persoonlijk goed met elkaar vinden. Openbare discussies, zoals die tussen randfiguur Brugmans en Kohnstamm op het Tweede Nederlandsch Pedagogisch Congres in 1930, waren tussen het vijftal onderling schaars. De inhoud van die discussie tussen Brugmans en Kohnstamm wijst de weg naar wat de academische pedagogen bond. Brugmans verdedigde een empirische pedagogiek, terwijl de academische pedagogiek van voor de Tweede Wereldoorlog juist werd gekenmerkt door een normatieve oriëntatie. Men deelde het inzicht dat een pedagogiek gedragen moest worden door een levensbeschouwing. En door inhoudelijke verschillen in levensbeschouwing heen bleek dat een stevige basis voor een goede professionele verstandhouding.

De leeftijdsverschillen binnen de groep waren aanzienlijk. De jongste, J. Waterink, geboren in 1890, scheelde twee generaties met nestor J.H. Gunning Wzn (1859). De welwillendheid van de laatste speelde vanzelfsprekend in die goede contacten een belangrijke rol. Na de aanvankelijke kritiek van Gunning op R. Casimir, de eerste collega waar hij mee te maken kreeg, groeide hun relatie uit tot wederzijdse bijstand en steun op allerlei gebied. De ontwikkeling van de verhouding tussen Gunning en Ph.A. Kohnstamm had een omgekeerd verloop. Aanvankelijk was de eenheid van gedachte erg groot, pas later werd de verhouding stroever, wat zich uitte in een verkapte kritiek op Kohnstamms *Persoonlijkheid in wording* in een bijdrage van Gunning op de jaarvergadering van de Vereniging voor Paedagogiek in 1929: "Over het begrip 'oefenen' in verband met karaktervorming". Goede persoonlijke verhoudingen stonden er borg voor dat een vinnige inhoudelijke discussie niet uitgroeide tot een echt conflict, al schijnt ook Casimir een belangrijke rol te hebben gespeeld in het bijleggen van de ruzie. Kohnstamms verweer en Gunnings repliek zijn uiteindelijk niet gepubliceerd.

Maar zelfs zo'n ingetogen conflict als dit tussen Gunning en Kohnstamm was uitzonderlijk onder de Nederlandse pedagogen. De verhoudingen varieerden van goed tot nog beter zoals van de misschien wat afstandelijke respectvolle relatie tussen Gunning en J. Hoogveld tot de wederzijds gevoelde waardering tussen Gunning en Waterink.

Wat voor soort wetenschapsgeschiedenis schrijft Mulder nu eigenlijk? Wat is het bijzondere van zijn aanpak? Wat onmiddellijk opvalt en aanspreekt is dat er uitzonderlijk veel aandacht besteed wordt aan de persoonlijke geschiedenis van de in vijf afzonderlijke hoofdstukken geschilderde figuren. Wat dat betreft kent het boek zijn weerga niet en de enige reden waarom je het niet in een ruk uitleest is dat het daar te dik voor is. Het boek stijgt ver uit boven de gebruikelijke opsomming van ideeën en probeert een directe relatie te leggen tussen persoonlijke levensgeschiedenissen en pedagogisch gedachtengoed. Mulder loopt zelfs het gevaar kritiek op te roepen dat hij zo opgaat in gepsychologiseer dat het boek als inleiding in de pedagogieken van de vooroorlogse universitaire pedagogen wel aan diepgang moet inboeten. Je kunt nu eenmaal niet alles.

Tegen die achtergrond lijken 250 pagina's tekst ineens weer aan de magere kant. Maar laten we eerst de door hem gekozen werkwijze als zodanig aan een kritische beschouwing onderwerpen. Is het eigenlijk wel mogelijk om op grond van het historisch bronnenmateriaal dat Mulder hanteert, van wetenschappelijke teksten tot brieven en autobiografieën, vast te stellen hoe de collega's werkelijk over elkaar dachten? Het huidige wetenschapsbedrijf, zoals ik het ken, wordt gekenmerkt door grote verschillen in onderlinge waardering. Als het om de in persoonlijke gesprekken geuite beoordelingen gaat is er een groot gebrek aan eenheid van gedachte over wat kwaliteit is en wie kwaliteit heeft. Dat lijkt ernstiger dan het is. Het is immers inherent aan een wetenschap die geen eenheid van methode kent. In de schriftelijke reacties die collega's op elkaars werk geven zijn die inconsistenties in beoordelingen nauwelijks terug te vinden, en hier en daar treft men zelfs een *esprit de corps* die gelijkenis vertoont met die van de vooroorlogse collega's, al is men vandaag de dag ook in persoonlijke relaties aanmerkelijk minder deftig dan toentertijd. Mulder moet het echter met die schriftelijke verslagen doen. Hoe mooi zijn "Onder professoren" ook zijn moge, het schiet, zelfs als serieuze variant, tekort omdat hij niet als W.F. Hermans deelgenoot en tijdgenoot was.

Maar in de manier waarop Mulder met zijn bronnen omspringt weet hij naar mijn sterke vermoeden die natuurlijke handicap van de historicus niet te compenseren. Ik merk dat hij uitspraken voor zoete koek slikt – zoals de opmerking van de zwager die Gunning typeerde door te spreken van "de pedanterie van iemand die zich zijn leven lang geroepen heeft gevoeld de menschheid te leeraren" (p. 71) – terwijl ik denk dat het om uitspraken gaat die elke amateur-psycholoog zou moeten wantrouwen. Ik merk dat hij in Gunnings beoordeling van Ligthart wel jaloezie als verklaring opvoert, terwijl ik daar in de betreffende stukken nu juist geen aanleiding voor kan ontdekken. Mijn bewondering voor Gunning is voor een deel juist gebaseerd op de ingetogen wijze waarop Gunning in de betreffende bespreking met de maar al te vaak irritant kwajongensachtige Ligthart omgaat. Mulder pikt net die ene zin eruit die het dramatisch hoogtepunt vormt van Gunnings kritiek op Ligthart, maar die in het bed van Gunnings instemming volstrekt anders

oogt dan in het bed van afkeuring dat Mulder ons voorschotelt. Voor de duidelijkheid: de kwalificatie 'Nederlands reformpedagogisch paradepaard', die Mulder ter aanduiding van Ligthart in die ene alinea over hun verhouding gebruikt (p. 87) is van Mulder en niet van Gunning.

Nu kan men natuurlijk stellen dat het hier om strijdige interpretaties gaat waarover te twisten valt. Daarover kan in dit ene voorbeeld niet uitputtend worden gesproken. Over de typering van de wijze waarop Mulder de grootvaders van de pedagogiek benadert kan echter geen misverstand bestaan. Hij neemt ze domweg niet echt serieus en verpakt zijn oordelen voortdurend in een wolk van misplaatste, quasi-minzame, onaantastbaarheid. Mulder is er niet in geslaagd om, binnen de mogelijkheden van de historicus, tijdgenoot te worden van personen die hij beschrijft en die merkwaardige afstand blijft zijn beschrijving het hele boek door parten spelen.

De vermoede willekeur komt ook tot uitdrukking in de manier waarop Mulder lijnen in leven en leer met elkaar probeert te verbinden. (We blijven onze voorbeelden nog even in het hoofdstuk over Gunning zoeken). Tegen de wijze waarop Mulder Gunnings visie op de rol van man en vrouw in de opvoeding en zijn pleidooi voor een vaderlijke opvoeding met diens eigen opvoeding verbindt kan men moeilijk bezwaar maken. Het is Gunning zelf die over dat verband geen enkel misverstand laat bestaan. Ook is het niet te wild om een relatie te veronderstellen tussen Gunnings latere pedagogische opvattingen over het belang van de bewaarschool en de buitenlandse leertijd van de jongelui en diens eigen jeugd. Maar naast nog meer interessante verbanden komt het volgende toch merkwaardig over. Mulder twijfelt aan Gunnings geheugen als deze in verband met de kern van zijn latere ethiek de tucht der vrijheid van het ouderlijk gezin roemt. Mulder vermoedt hier onjuiste weergave, interpretatie met terugwerkende kracht op basis van latere ethisch-religieuze ideeën of een combinatie van beide, en hij laat ons in het ongewisse over de oorsprong van zijn twijfel in dit geval en de oorsprong van zijn zekerheid in de andere gevallen. Dergelijke toevalligheden, naast het ontbreken van iedere methodische verantwoording van zijn 'psycho-historie' brengt mij in onzekerheid over hetgeen hij aan verbanden van deze soort te berde brengt.

Mulder zegt in zijn "Inleiding" nadrukkelijk geen internalistische geschiedschrijving van de Pedagogiek in Nederland tussen 1900 en 1940 te willen bieden. Dat externe factoren bij het ontstaan van de Pedagogiek een belangrijke rol hebben gespeeld laat zich raden, geen zinnig wetenschapshistoricus beperkt zich vandaag de dag tot de interne wetenschapsontwikkeling als het om de beschrijving van de ontwikkeling van een discipline gaat. Maar wat de oriëntatie op de specifieke culturele omstandigheden in de door Mulder beschreven periode betreft is de beschrijving veel te mager en te algemeen Europees om ook maar een fractie van die specifieke Nederlandse achtergronden te verduidelijken. Achteraf blijkt dat de beschrijving van de culturele achtergrond nog gratuit is ook, want ze figureert slechts plichtmatig op de eerste pagina's van het boek om later nergens echt terug te keren.

Omdat de eerste generatie universitaire pedagogen als pioniers een relatief nieuw vak opbouwden en niet of nauwelijks op Nederlandse voorgangers konden stemmen kiest Mulder voor een biografische opzet. Hij geeft zo een scherp inzicht in het individuele gesjouw en geploeter om de pedagogiek aan de universiteit gevestigd te krijgen. Maar het is de vraag of hij er in geslaagd is in zijn uitdrukkelijke bedoeling om *niet* een portrettengallerij samen te stellen. De voorgenomen beschrijving van de ontwikkeling van de universitaire pedagogiek als professionaliseringsproces komt er uiteindelijk bekaaid af. Het is de vraag of hij er in geslaagd is om voldoende licht te laten schijnen op de bovenpersoonlijke, sociale factoren die in de geschiedenis van de pedagogiek naar zijn zeggen een rol hebben gespeeld.

Een van de redenen waarom Mulder onzes inziens niet in die opzet slaagt is merkwaardigerwijs te vinden in het negeren van belangrijke interne factoren. Zijn beschrijving van de inhoud van diverse pedagogieken is zo oppervlakkig dat hij kansen laat liggen om door doordachte analyse van die inhoud het verband met externe factoren te verduidelijken. Neem nu de wijze waarop de pedagogen zelf hun vak definieerden. Daarin moet toch die strijd om het bestaan in gestolde vorm terug te vinden zijn. Daarin is niet alleen terug te vinden hoe men zich de discipline idealiter in haar verhouding tot praktijk en practici en in verhouding tot andere sociale wetenschappen en de wetenschap in het algemeen voorstelde, maar daarin moet men ook een entree kunnen vinden voor een adequate formulering van de vraag hoe die verhoudingen in werkelijkheid waren. Het werk van de besproken pedagogen blijft wat dat betreft betrekkelijk onaangeroerd. Daarmee laat de auteur een kans voorbij gaan werkelijk de diepte aan zijn verhaal te geven die hij voor ogen heeft. De vooroorlogse pedagogen besteedden stuk voor stuk veel aandacht aan wetenschapstheoretische kwesties, niet allen in de zelfde omvang maar wel met een zelfde intensiteit. Ik vind het zelfs jammer dat ze wat dat betreft sinds lang niet meer gelezen werden, omdat de wetenschapstheoretische discussies van de laatste vijfentwintig jaar best enige vakspecifieke inhoud konden gebruiken.

De discussie tussen Brugmans en Kohnstamm bijvoorbeeld is daarom interessant, omdat ze de gespannen verhouding tussen empirische psychologie en normatieve pedagogiek aan het licht kan brengen op een manier die ook voor onze tijd nog interessant is. Mulder beschrijft de discussie inhoudelijk slechts in algemene termen waardoor wel dui-

delijk wordt dat Kohnstamm superieuze argumenten moet hebben gehad, maar verborgen blijft welke deze argumenten nu precies waren. (Mulder excelleert overigens wel weer met een verwijzing naar het verslag van Gunning van de door Brugmans in Groningen afgenomen MO-examens, waarin te lezen stond dat deze "(...) vooral over psychologie examineerde en nauwelijks over pedologie en geschiedenis van de pedagogiek en dat hij een grote voorliefde voor indelingen, schema's en etiketten bleek te hebben" (p. 234), omdat hij daarmee zeer herkenbare achtergrondinformatie geeft.) Hetzelfde geldt voor het genoemde conflict tussen Gunning en Kohnstamm over gewoontevorming. Mulder geeft nauwelijks inzicht in de inhoudelijke discussie, maar des te meer weetjes over de persoonlijke verhoudingen.

Mulder concludeert dat er onder de pedagogen een betrekkelijk grote eenheid van gedachte bestond over de meeste centrale thema's zoals de verhouding van kind-zijn en volwassenheid, de opvoedingsnoodzaak, het onderscheid tussen opvoeding en socialisering, de rol van het gezin in de seksuele opvoeding, het belang van het goede voorbeeld en de zin van straffen, de verhouding van gezinsopvoeding en onderwijs, de waarde van reformpedagogisch gedachtengoed, enzovoort. Verschillen lijken volgens Mulder in het algemeen tot voor de hand liggende verschillen in persoonlijke achtergrond terug te voeren te zijn. Hij sluit zich op een gebruikelijke wijze bij het oordeel aan dat Kohnstamm als gunstige uitzondering gezien moet worden binnen een theoretisch wat onvoldragen pedagogiek en benadrukt de meer dan vreedzame coëxistentie tussen de pedagogen onderling, die vooral door de Vereniging ter Bevordering van de Studie der Pedagogiek en de MO-akte gestalte kreeg.

Mulders boek geeft een schat aan informatie. Het verschijnen van het werk vult een door velen gevoelde leemte op. Naast een aantal verspreide deelstudies ontbrak een geïntegreerd verhaal. Het werk doet in zekere zin wat het beschrijft, het emancipeert een emanciperende discipline. Identiteit heeft veel met geschiedenis te maken. Dat geldt zowel voor persoonlijke identiteit als voor de identiteit van een vak. In zo'n boek mogen personen en zakenregister niet ontbreken. Bij de produktie van het boek had men ook wat conscientieuzer tegen de tekstverwerkersduivel moeten strijden. De zogeheten harde afbreekstreepjes midden in de zinnen, de spatieloze zinnen, de nootnummers die over de letters heen vallen en de bladspiegel die nog juist binnen de pagina blijft, ze zullen binnen afzienbare tijd tot historische curiositeit geworden zijn.

Hoe het mogelijk is zo enthousiast te zijn over een boek waartegen men toch zoveel serieuze bezwaren heeft is mij een raadsel. Het moet iets van doen hebben met de fascinatie van het onderwerp. Die fascinatie wordt door het boek van Mulder alleen nog maar versterkt. Het gesignaleerde gebrek aan eenheid van gedachte over wat kwaliteit is lijkt in de pedagogiek echter manifester dan ooit. Mulder werd het doctoraat cum laude toegekend.

B. Levering

C. N. Brouwer, *Geïntegreerde lerarenopleiding; principes en effecten.* RUU, Utrecht, 1989, 512 pag., *f* 67,50, ISBN 90-9003172-3.

Het proefschrift van Brouwer is de neerslag van een evaluatie-onderzoek bij het PDI-RUU: de Utrechtse universitaire lerarenopleiding, in de tijd dat deze opleiding nog vier maanden besloeg. Een belangrijke component van de opleiding is het schoolpracticum (SP): een groeps-stage van drie aanstaande leraren, begeleid door een docent van de stageschool (de schoolpracticumdocent: SPD) en een docent van het opleidingsinstituut (de instituutsbegeleider: IB). Doel van het onderzoek was het beschrijven van (samenwerkingsverbanden in) het bestaande SP en verklaring van leereffecten die onder invloed daarvan optreden (p. 21). Een tweede onderzoeksdoel had betrekking op de gevolgen van het afschaffen van de taakuren voor SPD'en en leidde tot een afzonderlijke publikatie (N. Brouwer & S. Van den Heuvel, Roeien zonder riemen. Utrecht: RUU-IVLOS, 1987).

Het proefschrift van Brouwer bestaat uit drie delen. Het eerste deel heet 'Afbakening van het onderzoeksgebied'. Begonnen wordt met een concretisering van de onderzoeksdoelen in vijf onderzoeksvragen, betrekking hebbend op 1. de realisering van de opleidingsdoelen, 2. het tegengaan van de praktijkschok, 3. de samenwerking tussen opleidingsinstituut en stagescholen, 4. de honorering van de SPD en 5. aanbevelingen voor de inrichting van SP- en instituuts-perioden (hoofdstuk 1.)

Daarna (hoofdstuk 2) geeft Brouwer een kort overzicht van de geschiedenis van de universitaire lerarenopleiding, gevolgd door een presentatie van de drie opleidingsprincipes die constituerend zijn voor de Utrechtse opleidingsprogramma's, te weten: cyclische fasering van instituuts- en SP-perioden, begeleiding bij de individuele leerweg en intensieve samenwerking tussen opleiders. Vervolgens wordt een overzicht gepresenteerd van de Westduitse, Amerikaanse, Britse en Nederlandse onderzoeksliteratuur betreffende de praktijkschok van beginnende leraren. Het literatuuroverzicht wordt afgesloten met een door de auteur ontwikkeld onderzoeksmodel. In dat model is weergegeven langs welke lijnen de 'beroepscompetentie' van een leraar mogelijkerwijze wordt beïnvloed. Eén zo'n

lijn heeft betrekking op factoren in de beroepssituatie, de andere lijn heeft betrekking op opleidingsfactoren.

Het eerste deel van het proefschrift eindigt met een analyse van het begrip 'onderwijsbekwaamheid' als doel van delerarenopleiding(hoofdstuk 3). Er wordt een achttal clusters van vaardigheden onderscheiden. De auteur geeft per cluster een beschouwing aangaande doelen en inhouden die zijns inziens voor een lerarenopleiding van belang zijn.

Het tweede deel van het proefschrift van Brouwer, 'Uitvoering en resultaten van het onderzoek' start met een beschrijving van de onderzoeksopzet (hoofdstuk 4). Het betreft een longitudinaal onderzoek. Over een periode van ruim vier jaar werden 357 studenten, verdeeld over 24 opleidingen, gevolgd tijdens en na hun opleiding. Bij dit 'breedteonderzoek' werd voornamelijk gebruik gemaakt van kwantitatief te verwerken vragenlijsten (voor studenten, SPD'en en instituutsbegeleiders). De items van de vragenlijsten waren bedoeld als operationalisering van de componenten van het onderzoeksmodel. Daarnaast werd een 'diepteonderzoek' uitgevoerd bij een selectie van tien opleidingen. Per opleiding werden een instituutsbegeleider, twee SPD'en en een afgestudeerde met een baan in het VO geïnterviewd. Ook werden lesbezoeken afgelegd en stageverslagen bestudeerd.

De onderzoeksresultaten zijn gerapporteerd in twee hoofdstukken: 'De opleiding' (hoofdstuk 5) en 'De beroepspraktijk' (hoofdstuk 6). Elk van die hoofdstukken kent een 'beschrijvend deel' en een 'verklarend deel'. De beschrijvende delen geven een beeld van de gang van zaken in het SP en in de beroepspraktijk van de beginnende leraar. Ze zijn gebaseerd op gegevens uit het breedteonderzoek en het diepte-onderzoek. Ze bevatten veel citaten uit de interviews. In de verklarende delen wordt door middel van regressie-analyse nagegaan hoe de resultaten van de opleiding (startcompetentie en beroepscompetentie) samenhangen met de in het onderzoeksmodel opgenomen factoren. De startcompetentie zoals beoordeeld door de student zelf blijkt o.m. samen te hangen met de (beoordeling door de student van de leerzaamheid van de) afwisseling tussen instituuts- en SP-perioden. De beroepscompetentie van beginnende leraren blijkt vrijwel uitsluitend samen te hangen met factoren uit de beroepssituatie. Invloed van de opleiding blijkt via de regressie-analyse vrijwel niet geconstateerd te kunnen worden. Brouwer gaat voorts uitvoerig in op diverse verbanden tussen variabelen die de beroepssituatie van beginnende leraren betreffen. Een voorbeeld: "Zoals blijkt, voeren de beginnende leerkrachten bij wie spanningen van school in het privéleven doorwerkten, in hun werk beduidend meer verschillende aktiviteiten en werkzaamheden uit dan degenen die niet zulke negatieve effekten ondervonden" (p. 394).

In deel drie (hoofdstuk 7) van zijn proefschrift ('Slotsom') trekt Brouwer conclusies. Brouwer concludeert dat "een opleidingsprogrammering gericht op de integratie van theorie en praktijk volgens de drie opleidingsprincipes van het PDI-RUU (. . .) inderdaad (. . .) de startkompetentie van beginnende leerkrachten kan bevorderen" (p. 401). Tevens concludeert hij dat met betrekking tot het ontwikkelen van 'beroepscompetentie' de invloeden van de 'beroeps-socialisatielijn' overwegen over die van de 'opleidingslijn'. Vervolgens beantwoordt Brouwer zijn vijf onderzoeksvragen. Daartoe put hij uit de rijke verzameling data uit het breedte- en het diepteonderzoek. Ter beantwoording van onderzoeksvraag 5 worden aanbevelingen gedaan aan beleidmakers, opleiders, scholen en aanstaande en beginnende leerkrachten. Ter illustratie enkele aanbevelingen: "Verminder de klassegrootte in het voortgezet onderwijs tot 20" (p. 424). ". . . achten we het, in ieder geval zolang er geen officiële regelingen bestaan die het beginnende leerkrachten mogelijk maken zich in het kader van een volledige aanstelling geleidelijk in te werken (. . .), verstandiger niet meteen aan een volle baan te beginnen" (p. 532). Ten slotte worden ook aanbevelingen voor vervolgonderzoek gedaan.

Wat opvalt bij het lezen van het proefschrift van Brouwer is dat naast het genoemde onderzoeksdoel (beschrijving van SP en verklaring van leereffecten) impliciet een ander doel wordt nagestreefd, nl. vaststelling van de mate waarin en de condities waaronder de abituriënten van het PDI-RUU als beginnende leraar een praktijkschok doormaken. De modeltoetsing en de literatuurstudie over de praktijkschok hebben betrekking op dit impliciete doel. In relatie tot het wel genoemde onderzoeksdoel zou een literatuurstudie over inrichting van stages en leerprocessen in stages meer voor de hand hebben gelegen. Het deel van het onderzoek dat betrekking heeft op de praktijkschok is interessant. Naast een tamelijk uitputtend literatuuroverzicht biedt het proefschrift empirische gegevens over onder meer de invloed van klassegrootte, taakomvang en samenwerking met collega's. De citaten uit de interviews met de beginnende leraren geven een levendig beeld van hun situatie.

Het deel van het onderzoek dat betrekking heeft op het verklaren van leereffecten van de opleiding biedt minder. Een probleem dat Brouwer hier parten heeft gespeeld is dat de onafhankelijke variabelen niet erg varieerden. Immers alle respondenten handelden binnen het zelfde stage-model. Potentiële onafhankelijke variabelen met een grotere spreiding (bijvoorbeeld leerhoudingen en attitudes ten opzichte van onderwijs) werden niet in het onderzoek betrokken. Verwarrend is voorts dat evaluatief geformuleerde items in de regressie-vergelijkingen als opleidingskenmerken zijn ingevoerd. Een

oordeel over de lengte van de SP-perioden bijvoorbeeld zegt wellicht meer over de respondent dan over de feitelijke tijdsduur van de SP's. Ook de validiteit van de afhankelijke variabelen is discutabel: de student-oordelen over het effect van de opleiding correleren nauwelijks met de oordelen van de SPD'en. Voor de kritische lezer is het ontbreken van de schriftelijk afgenomen vragenlijsten in de bijlagen een handicap.

Eigenlijk zijn de meest interessante stukjes van dit deel van het onderzoek de twee variantie-analytische uitstapjes die Brouwer maakt naast zijn regressie-analytische studie: een vergelijking tussen inductief en deductief geprogrammeerde opleidingen (p. 273 e.v.) en een analyse van de invloed van de motivatie aan het begin van de studie op de resultaten van de opleiding (p. 275 e.v.).

Het proefschrift van Brouwer is een lijvig boekwerk. Brouwer heeft er voor gekozen alles wat hij te zeggen had op grond van zes jaar onderzoek in één band te rapporteren. Daardoor is het verslag minder 'strak' geworden dan wanneer de auteur had gekozen voor bijvoorbeeld alleen de verantwoording en rapportage van de model-toetsing. Het boek is echter voor de lerarenopleiding een bron waaruit veel informatie en ideeën kunnen worden geput. Het is te hopen dat het ook een impuls zal geven tot het verbeteren van de arbeidsomstandigheden van beginnende leraren. Het proefschrift van Brouwer laat (weer eens) zien dat daar alle reden toe is.

J. A. de Jong

W. A. W. de Graaf, *De zaaitijd bij uitnemendheid; jeugd en puberteit in Nederland 1900-1940,* Academisch Boeken Centrum, De Lier, z.j. (1989), 304 pag.

De geschiedschrijving over de sociale pedagogiek en over de jeugd in het algemeen beleeft in Nederland een opmerkelijke bloeiperiode. Weliswaar is de wetenschappelijke aandacht nog nagenoeg geheel op jongeren in de twintigste eeuw gericht, maar daar weten we nu dan toch inmiddels heel wat van. Verschillende monografieën zagen recentelijk het licht: over werkende jongeren (J. Peet, *Het uur van de arbeidersjeugd* (1987)); over de jeugdzorg in Amsterdam (M. Lunenberg, *Geluk door geestelijke groei*(1988)); over de jeugdzorg in Maastricht (P. P. Jansen e.a., *Van beschermelingen tot actief deelnemende jongeren* (1989)); over neo-calvinistische jongelings- en meisjesverenigingen (J. C. Sturm, *Een goede gereformeerde opvoeding* (1988)). En dan zie ik maar voorbij aan alle artikelen en doctoraalscripties op dit terrein.

In 1989 is hieraan toegevoegd de historische dissertatie van de Utrechtse ontwikkelingspsycholoog Willibrord de Graaf. Hij behandelt het ontstaan van de jeugd als aparte psychologische en pedagogische categorie tussen 1900 en 1940. Volgens De Graaf is er dan voor het eerst in de geschiedenis sprake van een duidelijk onderscheiden en problematische overgangsfase tussen kindertijd en volwassenheid, tussen gezin en arbeid. Het concept puberteit als een moeizame en conflictueuze fase in het leven van jongeren doet zijn intrede. De jongelui sluiten zich aaneen in jeugdbewegingen. Goedwillende ouderen werpen zich op als leiders van jeugdorganisaties. Het opgroeiende geslacht wordt in toenemende mate object van de bemoeiingen van pedagogen, psychologen en politici. De overheid zet zich tot het eerste van een eindeloze reeks rapporten over de problemen van jongelieden en jongedochters (commissie Van Wijnbergen, 1919). Psychologen beschrijven de kenmerken van puberteit en adolescentie. Pedagogen ontwerpen naast gezin en school een derde opvoedingsmilieu: de vrije tijd van jongeren.

Het nieuwe in de benadering van De Graaf is dat hij deze bekende gegevens in een tweeledig theoretisch kader plaatst. De ene betooglijn is dat de moderne puber in zekere zin geschapen is doordat al die wetenschappelijke onderzoekers en al die verontruste opvoeders ineens op een bepaalde manier gingen spreken en schrijven over de jeugd. Dit 'vertoog' (zoals De Graaf dat in navolging van zijn idool Foucault noemt) creëerde dus eigenlijk die warrige en woelige fase in het leven van de twintigste eeuwse tieners. De tweede betooglijn voert langs sociaal-economische ontwikkelingen na 1900: voortgaande mechanisatie in de industrie, rationalisatie van produktiewijzen, massificatie van de consumptie, en vooral ook de zware economische crisis van de jaren dertig. Beide factoren (het vertoog van deskundigen en de positie van jongeren in de arbeid) zorgen er volgens De Graaf voor dat jongeren na 1900 een nieuwe identiteit krijgen opgedrukt en voortaan een tijdje als puber door het leven moeten gaan.

In het eerste hoofstuk voert De Graaf de moderne jeugd na 1900 ten tonele, met alle psychologische kenmerken die men haar toedichtte en alle pedagogische en politieke maatregelen die men ging treffen om de gesignaleerde problemen te attaqueren. Daarbij werpt De Graaf de vraag op of die moderne jeugd nu inderdaad wel zo'n nieuw verschijnsel is in de geschiedenis. Hoewel hij een aantal interessante gegevens over jongeren vóór 1900 naar voren brengt, schort het hem toch duidelijk aan kennis van wat literatoren, filosofen, theologen en pedagogen al eeuwenlang met betrekking tot specifieke problemen en kenmerken van de jonkheid hebben gedebiteerd. Lezing van bijvoorbeeld Vadertje Cats of van de studies van de Vlaamse classicus-

historicus E. Eyben had hem ongetwijfeld genuanceerder kunnen doen denken over het door hem veronderstelde unieke karakter van het twintigste eeuwse jeugdleven.

In het tweede hoofdstuk bespreekt De Graaf de economische ontwikkelingen in de vroege twintigste eeuw, toegespitst op de vraag wat dit betekende voor de positie van jongeren in arbeid en onderwijs. Interessant is hier zijn gebruik van het onder economisch-historici momenteel in de belangstelling staande 'taylorisme' (rationalisatie, efficiëntie, stopwatch aan de lopende band) en het naar een bekende autofabrikant genoemde 'fordisme' (schaalvergroting, standaardisering van de produktie, creëren van koopkracht). Merkwaardig is overigens dat zowel de in het eerste hoofdstuk besproken visie op de jeugd, alsook de in het tweede hoofdstuk geschetste sociaal-economische veranderingen allebei welbeschouwd pas rond 1920 gedateerd moeten worden, waarmee de periodisering van De Graafs boek (1900-1940) feitelijk volstrekt op losse schroeven komt te staan.

In het derde hoofstuk reconstrueert De Graaf de ideeën en opvattingen met betrekking tot de relaties tussen arbeid en jongeren destijds. Daarbij komen verschillende interessante gegevens naar voren, bijvoorbeeld betreffende het sekseonderscheid in onderwijs en jeugdarbeid, verschillen tussen ambacht, huishouden en fabriek, de ontwikkeling van psychotechniek en beroepskeuzeadvisering en het nijverheidsonderwijs. In het algemeen mag gezegd worden dat het hier besproken boek een ongelofelijke informatie-dichtheid heeft. Talloze gegevens, verwijzingen, weetjes zijn verspreid in de tekst en in het zeer omvangrijke notenapparaat. Dit maakt het boek zowel waardevol, alsook onoverzichtelijk en nogal saai.

In hoofdstuk vier komen de specifieke psychologische en pedagogische opvattingen aan de beurt. Eerst geeft de auteur een beknopte schets van de vrije jeugdvorming (waarbij, zoals zo vaak voorkomt in de geschiedschrijving over jeugdorganisaties, de confessionele jeugdorganisaties er erg bekaaid afkomen). Dan komt een uitgebreide behandeling van toenmalige ideeën over gezin, school, arbeid en de ontwikkeling van jongeren daarbinnen. Dit hoofdstuk vind ik het meest geslaagde van het boek. Zeer bezwaarlijk is nochtans dat De Graaf in zijn beschrijving teveel probeert om een homogeen vertoog te (re)construeren, waardoor hij voorbijgaat aan de verzuiling en aan levensbeschouwelijke verschillen, die nu juist in de beleving van de door hem behandelde auteurs zo'n belangrijke rol speelden.

In hoofdstuk vijf laat De Graaf zien dat de economische crisis na 1929 en de daarmee gepaard gaande jeugdwerkloosheid versterkend en versnellend hebben gewerkt op de pedagogisering van (met name de arbeiders)jeugd in Nederland. De toewijding waarmee de schrijver op zoek is geweest naar betrouwbaar cijfermateriaal wekt bewondering op. Ook zijn analyse dat de jeugdopvoeders schijnbaar alle pubers tot hun cliëntèle rekenden, maar in feite voornamelijk de arbeidersjeugd bedienden, is alleszins de moeite waard.

In het slothoofdstuk zes vat de auteur zijn werk nog eens samen (bepaald geen overbodige luxe in dit boek!) en trekt hij een aantal prikkelende conclusies. Een belangstellende lezer met helaas weinig tijd, raad ik toch aan in ieder geval dit hoofdstuk eens door te nemen. Men dient wel te beschikken over doorzettingsvermogen en een milde houding ten aanzien van onzorgvuldig taalgebruik en slechte formuleringen.

J. C. Sturm

Mededelingen

Stipendia voor opleidingskundig onderzoek

Het bestuur van de Foundation for Corporate Education – FCE – Stichting Opleidingskunde stelt in 1990 een bedrag van *f* 45.000,— ter beschikking voor onderzoek op het terrein van opleidingen en leerprocessen in arbeidsorganisaties. Het bestuur nodigt onderzoekers uit om onderzoeksvoorstellen, die in aanmerking komen voor financiële ondersteuning in de vorm van stipendia, vóór 1 oktober 1990 in te dienen. Belangstellenden kunnen het stipendium-reglement opvragen bij de Stipendiumcommissie van de Foundation for Corporate Education, Dronkelaarseweg 13, 3784 WB Terschuur.

Najaarsconferentie Sectie Klinische- en Orthopedagogiek NVO

De sectie klinische- en orthopedagogiek van de NVO organiseert op 26 oktober 1990 te Utrecht een studiedag met als titel: 'Kind en gezin; gezinstaxaties in het diagnostisch proces'.

Voor nadere informatie en inschrijving kunt U zich wenden tot: NVO, Korte Elisabethstraat 11, 3511 JG Utrecht.

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
15e jaargang, nr. 3, 1990

Het biografisch perspectief in onderzoek van het professioneel handelen van leerkrachten, door C. Kelchtermans en R. Vandenberghe
De institutionalisering voorbij: over de duurzaamheid van vernieuwing in de school, door T. Jansen en R. van der Vegt
Een metacognitieve benadering van de doelstellingenproblematiek in het onderwijs, door L. de Klerk
Besprekingsartikel: Het project Onderwijs en Sociaal Milieu: een eindbalans, door M. Matthijssen

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 5, 1990

Computerondersteund onderwijs voor kinderen met ernstige leerproblemen: inleiding, door E.C.D.M. van Lieshout
Flash Cards; een vergelijking van Toepassingsprocedures, door K. van den Bosch, W.H.J. van Bon en R. Schreuder
De constructie van een computergestuurd programma voor aanvankelijk lezen, door A. van der Leij
Computergestuurd programma met betrekking tot de spellingregels van open en gesloten lettergrepen, door J.H.M. Hamers
Een gecomputeriseerde training voor het optellen en aftrekken tot 100 voor kinderen met rekenproblemen, door M.W.J. Baltussen en E.C.D.M. van Lieshout
Visuele modellen leren opbouwen in een computergestuurde training voor redactierekenen, door M.W.M. Japsers en E.C.D.M. van Lieshout
Computers voor kinderen met ernstige leerproblemen: een tussenstand, door P. Reitsma

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 6, 1990

Effecten van een sociaal-redzaamheidsprogramma voor zwakbegaafde jongeren (II), door J.K. Bleeker, D.T. van der Beek en H.T. van der Molen
Veranderend leraargedrag door middel van videoregistratie in de klas, door W. van Werkhoven en H. Brekelmans
Differentiatie en integratie binnen het speciaal onderwijs, door P.L.M. van der Doef en R.J. van der Krol

Ontvangen boeken

Bergeijk, J. van, W. van den Bon, A. Leeuwis (red.), *Agrarisch onderwijskundig onderzoek in ontwikkeling*. Landbouwuniversiteit Wageningen, 1990.

Carlier, I.V.E., *Overreding bij kinderen en jeugdigen* (dissertatie). Stichting Kinderstudies, Groningen, 1990, *f* 26,—.

Depaepe, M., *Meten om beter te weten?* Geschiedenis van de experimenteel-wetenschappelijke richting in de Westerse pedagogiek vanaf het eind van de 19e eeuw tot aan de Tweede Wereldoorlog (dissertatie). Katholieke Universiteit, Leuven, 1989.

Oenen, S. van, *Meesterschap en moederschap.* Teksten van Jan Ligthart en tijdgenoten als onderwerp van vrouwenstudies-pedagogiek (dissertatie). Acco, Amersfoort, 1990, *f* 32,—.

Absenteïsme in het voortgezet onderwijs*

K. TJ. BOS, A. M. RUIJTERS en
A. J. VISSCHER
Faculteit Toegepaste Onderwijskunde en het Onderzoek Centrum Toegepaste Onderwijskunde van de Universiteit Twente

Samenvatting

Ondanks de ruime aandacht die aan het spijbelprobleem wordt gegeven zijn de empirische gegevens omtrent de omvang van dit fenomeen beperkt. In dit artikel wordt gerapporteerd over een onderzoek naar de omvang van het geoorloofde en ongeoorloofde absenteïsme van leerlingen op zesendertig scholen in vier grote Nederlandse steden. Het verzuim werd op acht niveaus (onder andere per school, per leerjaar en per vak) gedurende drie schooldagen tijdens 8990 lessen vastgesteld. Uit de resultaten blijkt dat de omvang van het geoorloofd absenteïsme (4.7%) gemiddeld hoger ligt dan de omvang van het ongeoorloofd verzuim (spijbelen), nl. 4.4%. Tussen scholen zijn er aanzienlijke verschillen in de spijbelomvang. Op scholen met een lbo-afdeling wordt gemiddeld meer verzuimd. Dit geldt ook voor de meeste hogere leerjaren, de vrijdag, het eerste en het zevende tot en met tiende lesuur, de beroepsgerichte en de economische vakken. Bij vakken die veelal als 'spijbelvakken' beschouwd worden werden relatief lage verzuimpercentages gevonden. De meeste spijbelaars spijbelen maximaal twee lesuren per dag. Gemiddeld zestien procent van de leerlingen van een school spijbelt, hoewel dit percentage tussen de scholen sterk verschilt.

* Dit onderzoek is uitgevoerd in opdracht van het Ministerie van Onderwijs en Wetenschappen.

1 *Inleiding*

Sinds het begin van de tachtiger jaren wordt in Nederland veel aandacht besteed aan de problematiek van het verzuim van leerlingen in het voortgezet onderwijs. In 1986 bracht de toenmalige Staatssecretaris van Onderwijs en Wetenschappen de beleidsnotitie 'Schoolverzuim' uit (Ginjaar-Maas, 1986). Daarin wordt onder meer zorg uitgesproken over aard en omvang van het schoolverzuim in het voortgezet onderwijs.

Voor een goed inzicht in de omvang en aard van het verzuimprobleem is het van groot belang om schoolverzuim *eenduidig* te definiëren en de omvang van het schoolverzuim op een *exacte* manier te meten. In dit artikel worden de resultaten beschreven van een studie waarin het verzuim op deze wijze bestudeerd is.

2 *Definitie en meting van schoolverzuim*

In verschillende studies worden vaak uiteenlopende definities van de omvang van het schoolverzuim (= spijbelen) gehanteerd. Dit heeft gevolgen voor de vergelijkbaarheid van de uitkomsten van die studies.

De definities van schoolverzuim variëren van 'wat de school eronder verstaat' (Bos, Ruijters & Visscher, 1989) tot 'schoolverzuim is *alle* afwezigheid van leerlingen tijdens lesuren waarop ze wel aanwezig zouden moeten zijn' (De Vries, Monsma & Mellink, 1989). Geilman (1984) definieert de verzuimomvang als het aantal leerlingen dat door de school als spijbelaar is geregistreerd. In zijn onderzoek is het aantal verzuimers op één dag in januari vastgesteld door middel van een vergelijking tussen een steekproef uit de administratie van de leerplichtige leerlingen van de gemeente Rotterdam en het daadwerkelijke bezoek van deze leerlingen aan hun school. Geconcludeerd wordt dat 1.4% van de leerlingen spijbelt. 't Hoen, Bosch en Verhagen (1986) deden hetzelfde voor de stad Arnhem.

In een onderzoek van De Vries et al. (1989) naar de invloed van de school op schoolver-

zuim wordt een erg ruime verzuimmaat gehanteerd. De omvang van het schoolverzuim wordt hierin opgevat als alle afwezigheid van leerlingen, met uitzondering van afwezigheid in verband met langdurige ziekte. Hierbij geldt dat ongeacht het aantal lesuren waarop een leerling op een dag afwezig is geweest de omvang van het verzuim voor die dag als één volledige dag wordt geregistreerd. Als reden hiervoor geven de onderzoekers "...op veel scholen wordt de afwezigheid op uren later op de dag minder goed bijgehouden en het risico is daardoor groot dat juist die scholen minder verzuim zouden lijken te hebben dan feitelijk het geval is" (pag. 3).

Met deze verzuimmaat wordt het werkelijke verzuim systematisch overschat hetgeen de onderzoekers zelf ook toegeven. Zij zien echter als groot voordeel van deze maat dat ze compenseert voor zogenaamd verborgen verzuim (afwezigheid onder het mom van smoezen die niet door de school zijn te controleren). De resultaten van deze studie, die in het leerjaar havo-4 van zes verschillende scholengemeenschappen voor havo/vwo werd uitgevoerd, wijzen uit dat gemiddeld 11% van deze leerlingen één of meerdere uren per dag afwezig is.

In ander Nederlands onderzoek is aan leerlingen of de schoolleiding gevraagd een schatting te geven van de omvang (en duur) van het verzuim. Jaspers, Schut, De Keyser, Gunnink en Hüsken (1984) hebben aan leerplichtige leerlingen in Leeuwarden gevraagd hoeveel uren ze in een bepaalde periode spijbelden. Dorsteen (1985) deed hetzelfde bij leerlingen die naar het TV-programma 'Je ziet maar' keken. Dit waren vooral leerlingen uit de Randstad. Uit beide studies kan volgens de onderzoekers worden geconcludeerd dat ongeveer 85% van alle leerlingen 'nooit' of 'bijna nooit' spijbelt. De overige 15% van de leerlingen spijbelt 'wel eens' of vaker. In het onderzoek van Jaspers et al. (1984) is gebleken dat van deze spijbelaars driekwart maandelijks, 20% wekelijks en 5% dagelijks spijbelt. Uit dit onderzoek kan worden afgeleid dat de leerlingen gezamenlijk 1,5% van alle schooldagen spijbelden (De Vries & Peetsma, 1987).

In een studie van De Vries en Peetsma (1987) werd aan schoolleiders gevraagd hoe groot naar hun mening de omvang van het spijbelen is. Hun oordelen vallen beduidend lager uit dan die van de leerlingen uit het eerdergenoemde onderzoek van Jaspers et al. De verschillen tussen de schattingen van de verzuimomvang door de schoolleiding enerzijds en leerlingen anderzijds kunnen waarschijnlijk gedeeltelijk worden toegeschreven aan het volgende. Leerlingen nemen in de schatting van hun eigen verzuim ook de door hun ouders of verzorgers gedekte smoezen mee, terwijl schoolleiders dit bij hun schatting uiteraard niet kunnen doen. De laatsten baseren hun schattingen van de omvang van het schoolverzuim op de verzuimregistratie van hun school. Deze verklaring wordt door onderzoek ondersteund; volgens Jaspers et al. (1984) registreren scholen slechts één van de acht spijbelaars. Zij komen tot dit verhoudingsgetal door de percepties van de leerlingen met die van de schoolleiding te vergelijken. Ook het onderzoek van Bos, Van Kesteren, Stoel en Vermeulen (1990) wees dit uit. In dit onderzoek werd onder meer aan leerlingen gevraagd hoe zij over de verzuimregistratie van hun school denken. De leerlingen meldden dat de registratie per docent verschilt en dat zij het idee hadden dat zij, als zij dat zouden willen 'ongemerkt' één of meerdere uren weg zouden kunnen blijven. Onderzoek naar de omvang van schoolverzuim gebaseerd op gegevens uit de schooladministratie (zie Lesman, 1980; Vlug, 1985; Grimshaw & Pratt, 1986) heeft met andere woorden de nodige beperkingen.

In het buitenland is het beeld niet fundamenteel anders. Definities lopen uiteen (Galloway, 1985; Grimshaw & Pratt, 1986). Reid (1985) wijst erop dat de exacte betekenis van een verzuimpercentage nogal eens onduidelijk is. Onderzoeksresultaten zijn door de verschillende conceptualiseringen en dataverzamelingsmethoden niet of nauwelijks vergelijkbaar.

Er moet geconcludeerd worden dat de tot nu toe verrichte studies een aantal belangrijke tekortkomingen kennen en dat de kennis omtrent de omvang van het schoolverzuim beperkt is. De onderzoeksresultaten zijn moeilijk vergelijkbaar door verschillen m.b.t.:

- de definitie van het begrip;
- de wijze van dataverzameling (leerlingpercepties, schoolleidingpercepties, gegevens uit de schooladministratie);
- de gehanteerde verzuimmaten (o.a. categorieën als 'wel eens'/'regelmatig'/'vaak',

een percentage van het totale aantal lesuren, of 'meer dan 4 uren per week').

Daarnaast is de empirische basis voor algemene uitspraken over de verzuimomvang niet groot: verzuimgegevens over één schooldag (Geilman, 1984; 't Hoen e.a., 1986), enkele scholen van één schooltype (Lesman, 1980), één stad (Jaspers e.a., 1984; Geilman, 1984; 't Hoen e.a., 1986).

De vaak gebruikte verzuimschattingen van leerlingen en leerkrachten en de gegevens uit de schooladministratie over- of onderschatten de werkelijke verzuimomvang. De Vries en De Jong (1986) geven aan dat de schattingen zeer uiteenlopen, mede doordat leerlingen en schoolleiders, die deze schattingen geven, niet hetzelfde onder spijbelen verstaan. Hierboven werd reeds aangegeven dat het nadeel van de schooladministratie is dat deze slechts een klein deel van de werkelijke absenties bevat.

Het is blijkbaar erg moeilijk om schoolverzuim nauwkeurig vast te stellen. Dat neemt echter niet weg dat gestreefd moet worden naar zo exact mogelijke gegevens. We denken daarbij aan verzuimgegevens die zo nauwkeurig mogelijk (liefst numeriek) de verzuimomvang weergeven. Een heldere definitie van wat onder verzuim verstaan wordt en met name hoe de verzuimmaat berekend wordt zijn daarbij gewenst. Wat de dataverzameling aangaat, dienen ongeoorloofde en geoorloofde absentie van elkaar onderscheiden te worden. De reden hiervoor is dat sommige leerlingen waarschijnlijk onder de vlag van 'geoorloofde absentie' spijbelen (Bos et al., 1990). Bovendien is het van belang dat de verzuimomvang gerelateerd wordt aan verschillende schooltypen, leerjaren, lesuren, schooldagen en vakken, omdat de verzuimomvang op deze niveaus kan verschillen.

Gezien het belang van een eenduidige definiëring van de kernbegrippen geven we hier nu onze definities weer.

Onder *schoolverzuim* (ongeoorloofde absentie) verstaan wij: het zonder een door de school als geldig beschouwde reden afwezig zijn van leerlingen tijdens een lesuur. *Geoorloofde absentie* wordt hier gedefinieerd als het met een door de school als geldig beschouwde reden afwezig zijn van leerlingen tijdens een lesuur.

Deze definities hebben hun beperkingen. Idealiter zouden onderzoekers zelf criteria voor geoorloofde absentie moeten vaststellen en van elke geconstateerde absentie moeten nagaan of deze als geoorloofd of ongeoorloofd beschouwd moet worden. Aangezien dit in de praktijk niet haalbaar is, bepaalt elke school zelf of een absentie al of niet geoorloofd is. Scholen verschillen echter ten aanzien van wat als een geoorloofde absentie wordt gezien. Waarschijnlijk zijn er in dit opzicht zelfs binnen scholen nog verschillen. Bovendien is de ene school nauwkeuriger dan de andere in het afhandelen van absenties (d.w.z. het bepalen of het geoorloofde of ongeoorloofde absentie betreft).

Een ander probleem betreft het feit dat het percentage geoorloofde absenties waarschijnlijk ook ongeoorloofde absenties bevat, omdat leerlingen en/of ouders smoezen geven die door de school geaccepteerd (moeten) worden. Deze beperkingen zullen onderzoekers moeten accepteren als men niet over de middelen beschikt om zelf van elke absentie de reden te achterhalen.

De verzuim*omvang* kan per niveau worden uitgedrukt in een *verzuimpercentage*. Dit is het percentage gespijbelde lesuren van het totaal aantal lesuren dat alle leerlingen per niveau waarop het verzuim is berekend aanwezig zouden moeten zijn. Voorbeelden van dergelijke niveaus zijn: de school, de schooltypencombinatie, het schooltypeleerjaar, het leerjaar, de dag, het lesuur en het vak.

De *omvang van het verzuim* wordt gedefinieerd als het verzuimpercentage per niveau. Vaststelling van de omvang van het schoolverzuim in de scholen levert slechts een deel van de informatie op die nodig is om inzicht te krijgen in het verzuimprobleem. Een verzuimpercentage zegt iets over de mate waarin lessen gemiddeld worden verzuimd door *alle* leerlingen gezamenlijk, maar geeft nog geen inzicht in de proportie leerlingen die spijbelt. Evenmin geeft het inzicht in het gemiddelde aantal gespijbelde uren *per spijbelaar* in een bepaalde periode. In verband hiermee is in dit onderzoek per school eveneens het percentage leerlingen berekend dat op de meetdagen één of meerdere lesuren heeft gespijbeld, evenals het percentage spijbelaars dat een bepaald aantal lesuren per meetdag heeft gespijbeld.

Terwijl de omvang van het schoolverzuim aangeeft door *hoeveel* leerlingen van een school er gemiddeld per lesuur op de meetdagen gespijbeld wordt, kan het percentage leer-

lingen dat hiervoor verantwoordelijk is uitgedrukt worden in de *spijbelaarsratio*. Dat is de verhouding tussen het aantal verschillende spijbelaars op de meetdagen en het totale aantal leerlingen dat bij de verzuimmeting betrokken was.

3 *Vraagstellingen*

In dit artikel wordt getracht antwoord te geven op drie vragen:

1. Hoe groot is de verzuimomvang in scholen voor voortgezet onderwijs in vier grote Nederlandse steden?
2. Hoe groot is de gemiddelde verzuimomvang per schooltypencombinatie, schooltypeleerjaar, leerjaar, dag, lesuur, lesuur per dag en vak?

3a. Hoe groot is het gemiddelde percentage *leerlingen* dat op de drie meetdagen 1 of meer lesuren heeft gespijbeld?

3b. Hoe groot is per meetdag het gemiddelde percentage *spijbelaars* dat respectievelijk 1 of 2 lesuren, 3, 4 of 5 lesuren, of meer dan 5 lesuren heeft gespijbeld?

4 *Methode*

4.1 *Onderzoeksgroep*

De onderzoeksgroep werd gevormd door 36 scholen uit Haarlem, Amsterdam, Rotterdam en Utrecht die deelnamen aan een project gericht op de terugdringing van het schoolverzuim met behulp van een computerondersteund absentieregistratiesysteem. Uit Figuur 1 blijkt dat tweederde van de scholen relatief klein (tot 500 leerlingen) is en dat de lbo/lbo-scholengemeenschappen een vrij groot gedeelte van de totale groep uitmaken (10 van de 36). Gegeven de selectheid van de onderzoeksgroep is het niet mogelijk om de resultaten van deze studie te generaliseren.

4.2 *Procedure en instrumenten*

De verzuimmeting heeft op alle scholen in alle leerjaren (behalve in het examenleerjaar) op maandag, woensdag en vrijdag van één week in april 1988 plaatsgevonden. Aangezien de verzuimomvang kan variëren tijdens het schooljaar (Karweit, 1973; Baum, 1978; Billington, 1978) kan op basis van de verzamelde gegevens geen uitspraak worden gedaan over

Schoolgrootte	0-300 lln.	300-500 lln.	500-750 lln.	750 en meer lln.	totaal
Schooltypencombinatie	aantal	aantal	aantal	aantal	aantal
Categoriale MAVO	6	1	-	-	7
Categoriale LBO of IBO	3	1	1	-	5
Scholengem. IBO-LBO	-	7	2	1	10
Scholengem. LBO-MAVO	1	1	1	-	3
Scholengem. LBO-AVO	-	-	-	2	2
Scholengem. AVO-VWO	-	-	3	3	6
Overig (middenschool 2x ISK 1x)	2	1	-	-	3
Totaal	12	11	7	6	36

Figuur 1 *Aantallen projectscholen gegroepeerd naar groepsgrootte en schooltypen (combinaties)*

de overige delen van het schooljaar.

De docenten van de scholen vulden tijdens alle lessen die zij op de drie meetdagen gaven een standaardformulier (de absentielijst) in. Hierop werden achtereenvolgens voor elke les die een docent gaf genoteerd: het lesuur, het vak, het aantal leerlingen dat aanwezig behoort te zijn tijdens die les, het schooltype en het leerjaar van de groep waaraan les wordt gegeven, evenals de namen van de afwezige leerlingen. Deze data vormden de basis voor de berekening van de verzuimpercentages.

De absentie-afhandelaar van elke school (de persoon die nagaat wat de reden van de afwezigheid van elke absente leerling is en naar aanleiding daarvan eventueel actie onderneemt) vulde voor elke meetdag ook een standaardformulier (de lijst met geoorloofde absenten) in. Hierop vermeldde hij of zij achtereenvolgens: de naam van de *geoorloofd* absente leerling, het schooltype en leerjaar van de leerling, alsmede het aantal uren dat de betreffende leerling op die meetdag geoorloofd afwezig is geweest. Door de absentielijsten van de docenten en de door de absentieafhandelaars ingevulde lijsten met geoorloofde absenten met elkaar te vergelijken kon per meetdag het aantal *ongeoorloofde* absente leerlingen worden vastgesteld.

Omdat de namen van de ongeoorloofd absente leerlingen bekend waren kon per ongeoorloofd absente leerling ook het aantal gespijbelde uren per meetdag worden bepaald. Deze informatie is gebruikt om per school het percentage leerlingen vast te stellen dat op de drie meetdagen één of meer lesuren heeft gespijbeld (vraag 3a) en het percentage spijbelaars dat een bepaald aantal lesuren per meetdag spijbelde (vraag 3b). Het was op basis van de beschikbare data niet mogelijk de laatstgenoemde percentages uit te splitsen naar vak, schooltype, lesuur etcetera, hetgeen voor de verzuimpercentages wel mogelijk was.

4.3 *Analyses*

De data zijn geanalyseerd door middel van beschrijvende statistische grootheden: gemiddelden, standaarddeviaties en ranges.

5 *Resultaten*

5.1 *Respons*

Alle 36 projectscholen hebben aan de meting deelgenomen. Het responspercentage van de absentielijsten is het percentage lessen waarvoor de betreffende docenten de absentielijst hebben ingevuld van het totaal aantal lessen dat volgens het actuele lesrooster door hen zou moeten zijn gegeven. Dit getal is gecorrigeerd voor ziekte van docenten gedurende de meetdagen. Het totale aantal lesuren, waarover verzuimdata zijn verzameld bedraagt 8990. Het responspercentage is per school berekend. Voor de projectscholen bedraagt het gemiddelde responspercentage 82%. Deze respons is zeer bevredigend te noemen. Bij de in Figuur 1 genoemde schooltypen ligt de standaarddeviatie van het responspercentage tussen de 8% en 12%.

5.2 *Verzuimomvang*

Alvorens we de verzuimpercentages bespreken willen we eerst enige algemene opmerkingen maken over de resultaten van de verzuimmeting. Per niveau zijn gemiddelden, standaarddeviaties en ranges berekend.

Hoe moet een (gemiddeld) verzuimpercentage van bijvoorbeeld 5% in dit onderzoek nu worden geïnterpreteerd? Een verzuimpercentage van 5% wil zeggen dat gemiddeld 1 op de 20 leerlingen bij een willekeurige les of lessen (dit is afhankelijk van het niveau waarop het verzuimpercentage is bepaald) ongeoorloofd afwezig is (spijbelt), terwijl deze leerling wel aanwezig zou moeten zijn. Het aantal *verschillende* leerlingen dat spijbelde is niet uit het verzuimpercentage af te leiden. Daarom is per school ook de spijbelaarsratio berekend; hierop wordt later ingegaan.

Verzuimomvang op schoolniveau

De verzuimomvang is per school over alle meetdagen vastgesteld. Het gemiddelde verzuimpercentage per school voor de gehele week bedraagt 4.4% (zie Tabel 1).

Dit houdt in dat over de drie meetdagen tesamen gemiddeld per school 4.4% van de potentiële spijbelaars (dit is het aantal leerlingen dat tijdens de gegeven lesuren aanwezig had moeten zijn) werkelijk spijbelde. Het verzuimpercentage varieert voor 35 scholen (van één school waren de verzuimdata te laat binnen) van 0.2% tot 13.8%. De standaarddevia-

Tabel 1 *Gemiddelde (M), standaarddeviatie (sd) en range (N=35) per school: verzuimpercentage en percentage geoorloofde absentie en totaal absentiepercentage*

Variabelen	M %	sd %	range %-%
Verzuimpercentage	4.4	3.8	0.2-13.8
Geoorloofd absentiepercentage	4.7	1.8	1.6- 8.4
Totaal absentiepercentage	9.1	4.3	2.5-18.2

tie bedraagt 3.8%. Deze resultaten wijzen op grote verschillen in de verzuimomvang tussen de scholen.

Eerder is betoogd dat vergelijking tussen de resultaten van dit onderzoek en die van andere studies problematisch is. De hier gehanteerde methode komt nog het meest overeen met die uit het onderzoek van Geilman (1984). In zijn onderzoek werd niet de afwezigheid per lesuur geregistreerd, maar werd gepeild welke leerlingen op één bepaalde dag ongeoorloofd afwezig waren. Zogenaamde 'uurtjes-pikkers' worden in de verzuimmaat van Geilman niet en in de studie waarover hier gerapporteerd wordt wel meegenomen. Dit zou het verschil in resultaat (Geilman vond een percentage spijbelaars van 1.4%) met de hier gepresenteerde gegevens kunnen verklaren.

Uit Tabel 1 blijkt verder dat het gemiddelde *geoorloofde* absentiepercentage 4.7% bedraagt (standaarddeviatie 1.8%). Dit is hoger dan het gemiddelde percentage ongeoorloofde absentie. Dit wijst erop dat op veel scholen relatief veel leerlingen in verband met ziekte, tandartsbezoek, familie omstandigheden e.d. absent zijn. In hoeverre dit altijd de ware redenen van hun afwezigheid zijn is niet bekend.

De som van het gemiddelde percentage verzuim èn geoorloofde absentie bedraagt 9.1%. Op basis van dit percentage kan men zeggen dat in de in dit onderzoek onderzochte steden van de honderd leerlingen per lesuur er gemiddeld negen om uiteenlopende (geoorloofde en ongeoorloofde) redenen afwezig zijn.

Verzuimomvang per schooltypencombinatie

Op *school*niveau is de verzuimomvang per schooltypencombinatie berekend. Hierbij is de combinatie van de schooltypen waaruit elke school is opgebouwd het uitgangspunt.

Tabel 2 *Verzuimpercentage per schooltypencombinatie op schoolniveau. Gemiddelde (M), standaarddeviatie (sd), range en aantal scholen (N)*

schooltypencombinatie op schoolniveau	M %	sd %	range %-%	N
LBO/LBO (- sgm.)	5.9	3.5	0.5-12.1	14
LBO/AVO (- sgm.)	6.6	5.6	1.4-13.8	6
AVO/VWO (- sgm.)	2.0	1.5	0.2- 5.1	16

Elk van de genoemde combinaties bevat zowel scholengemeenschappen als categoriale scholen. De verschillen in gemiddelde verzuimomvang tussen lbo/lbo- en lbo/avo-scholen enerzijds (5.9% en 6.6%) en avo/vwo-scholen anderzijds (2.0%) zijn aanzienlijk, evenals de verschillen in de spreiding. Een verklaring voor de verschillen in gemiddelde verzuimomvang zou kunnen zijn dat de leerlingenpopulatie op scholen met een lbo-afdeling meer problemen heeft dan die op avo/vwo-scholen en dat ze daarom spijbelgevoeliger is. Uit eerder onderzoek is bekend dat de sociaal-economische achtergrond (huisvesting, gezinsinkomen) van lbo-leerlingen zwakker is dan die van avo-leerlingen (o.a. Stoel, 1986).

Verzuimomvang per schooltypeleerjaar

De gemiddelde verzuimomvang verschilt niet alleen tussen schooltypencombinaties, maar ook tussen de onderscheiden schooltypeleerjaren. Op basis van de beschikbare verzuimdata bleken veel verschillende combinaties van een schooltype met een leerjaar mogelijk te zijn. Een indikking was derhalve noodzakelijk. De criteria waarop deze is gebaseerd zijn: het aantal beschikbare cases (per schooltypeleerjaar) en de relevantie van het schooltypeleerjaar voor het onderwijsveld. Na de indikking bleven 16 verschillende betekenisvolle schooltypeleerjaren over.

Bij vergelijking van de gemiddelde verzuimomvang van de verschillende schooltypeleerjaren valt bij bijna alle schooltypen een toename in de spijbelomvang met het leerjaar, te constateren. Voor het ibo, het lbo, het havo en het vwo is deze toename relatief groot, hoewel hierbij opgemerkt moet worden dat de standaarddeviatie van de verzuimpercentages

Tabel 3 *Verzuimpercentage per schooltypeleerjaar. Gemiddelde (M) standaarddeviatie (sd), range en aantal scholen (N)*

Schooltypeleerjaar	M %	sd %	range %-%	N
Brug - 1	2.5	3.5	0-16.9	33
Brug - 2	3.9	4.2	0-12.2	14
Brug - 3	6.0	8.9	0-19.0	5
IBO - 2	5.6	5.4	0-15.2	8
IBO - 3	9.4	10.1	0-33.3	11
LBO - 2	5.5	3.2	0-10.5	13
LBO - 3	8.1	6.4	0.5-23.1	17
MAVO - 2	1.6	1.4	0- 4.2	11
MAVO - 3	3.7	5.8	0.1-22.8	15
HAVO - 2	2.6	1.9	0.4- 3.9	3
HAVO - 3	1.7	2.3	0- 5.8	7
HAVO - 4	7.0	6.6	0.7-20.7	7
VWO - 2	0.5	0.9	0- 1.8	4
VWO - 3	2.3	2.1	0- 5.0	7
VWO - 4	1.9	2.2	0- 6.5	7
VWO - 5	5.7	5.6	0-14.4	7

van de ibo-leerjaren eveneens relatief hoog is te noemen. Een uitzondering wordt gevormd door het havo en vwo waarin een daling in de verzuimomvang te zien is in respectievelijk het derde en het vierde leerjaar. Een hypothese op basis van deze gegevens zou kunnen zijn dat op lagere vormen van voortgezet onderwijs de leerlingen op jongere leeftijd gaan spijbelen.

In de leerjaren vóór het examenleerjaar is de verzuimomvang in alle schooltypen het hoogst. Met name vallen hierbij de sterke toename in het verzuimpercentage van havo-3 naar havo-4 en die van vwo-4 naar vwo-5 op. Deze bevinding stemt overeen met de resultaten van eerdere studies (Babeliowsky, 1986; Jaspers e.a., 1984).

In Tabel 2 bleek dat het gemiddelde verzuimpercentage voor lbo/avo-scholen relatief hoog is (6.6%). Uit dit resultaat gecombineerd met de resultaten van Tabel 3 kan nu worden geconcludeerd dat dit hoge percentage waarschijnlijk vooral toegeschreven moet worden aan de lbo-afdelingen van die scholen.

Verzuimomvang per dag

De verzuimmeting heeft op drie dagen van één week plaatsgevonden: maandag, woensdag en vrijdag. Om te zien of de omvang van het spijbelen per dag verschilt zijn voor elk van deze drie dagen verzuimpercentages berekend.

De gemiddelde verzuimpercentages voor maandag (4.0%) en woensdag (4.2%) zijn ongeveer 1% lager dan het gemiddelde verzuimpercentage op vrijdag (5.0%). Dit gegeven is in overeenstemming met ander onderzoek (Karweit, 1973; Baum, 1978 & Billington, 1978), waaruit bleek dat met name op vrijdag meer gespijbeld wordt. Een vergelijking van het gemiddelde verzuimpercentage op schoolniveau met de gemiddelde verzuimpercentages per dag leert dat de verzuimomvang op vrijdag hoger ligt dan de gemiddelde verzuimomvang per school. Op maandag en woensdag is dit juist andersom. De standaarddeviatie van de verzuimpercentages van de 35 scholen op vrijdag (4.6%) is groter dan die voor de maandag (3.7%) en woensdag (3.5%). Deze grotere spreiding van verzuimpercentages op vrijdag wijst erop dat op een aantal scholen het verzuim op de laatste dag van de week grotere vormen aanneemt dan op de beide andere meetdagen, terwijl dit op andere scholen veel minder het geval hoeft te zijn.

Verzuimomvang per lesuur

De omvang van het verzuim verschilt per lesuur. Uit de gegevens blijkt dat het gemiddelde verzuimpercentage tijdens het eerste lesuur (5.2%) duidelijk hoger is dan tijdens de andere ochtenduren. Op het tweede, derde en vierde lesuur ligt het verzuimpercentage rond de 4%. Hierbij dient bedacht te worden dat sommige scholen te laat komen als spijbelen registreren en dat dit vooral voor het eerste lesuur gedaan wordt. Vanaf het vijfde lesuur is het gemiddelde verzuimpercentage hoger dan 4%: 4.7%, 4.5%, 5.6%, 7.7%, 12.1% en 12.8% voor respectievelijk het 5e, 6e, 7e, 8e, 9e en 10e lesuur. Het verzuim neemt vanaf het 7e lesuur aanmerkelijk toe (zie Figuur 2).

De gemiddelde verzuimpercentages van de lesuren negen en tien zijn in vergelijking met de andere percentages erg hoog. Hierbij moet het geringe aantal scholen dat een negende en tiende lesuur op het lesrooster heeft staan (respectievelijk 11 en 2 scholen) in ogenschouw

worden genomen. Bij het eerste tot en met 7e lesuur waren daarentegen steeds 34 of 35 scholen betrokken, bij het 8e lesuur 28.

Verzuimomvang per lesuur per dag

De omvang van het verzuim is per school niet alleen per lesuur (het eerste, tweede, derde etcetera lesuur) en per dag (op maandag, woensdag en vrijdag), maar ook per lesuur per dag (bijvoorbeeld maandag het derde lesuur of vrijdag het zevende lesuur) berekend.

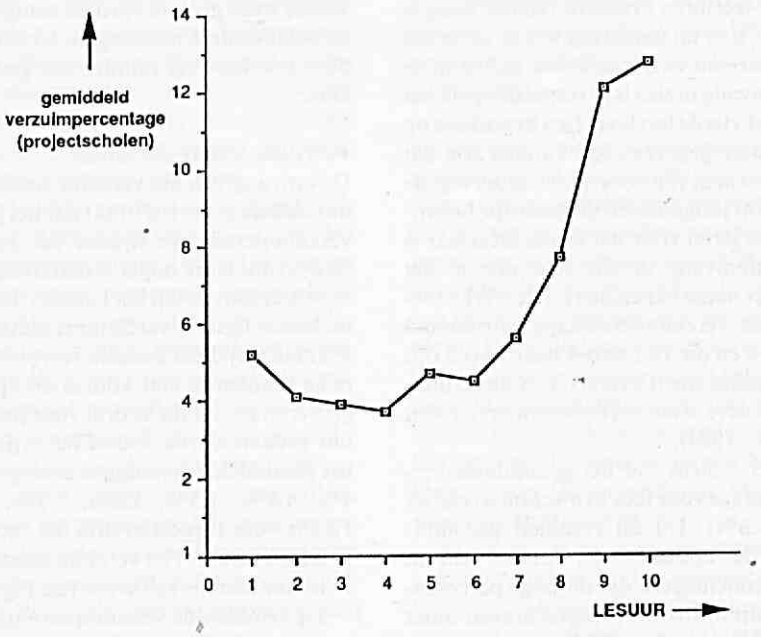
In verband met de beperkte beschikbare ruimte zullen we hier niet de gegevens voor alle dertig lesuren weergeven, maar een aantal opvallende resultaten bespreken.

Het gemiddelde verzuimpercentage op de eerste zeven lesuren op vrijdag is hoger dan dat op maandag en woensdag. Eerder zagen we dat het gemiddelde verzuimpercentage op vrijdag hoger is dan op maandag en op woensdag. We weten nu dat dit samenhangt met de hogere verzuimomvang tijdens de eerste zeven lesuren op vrijdag.

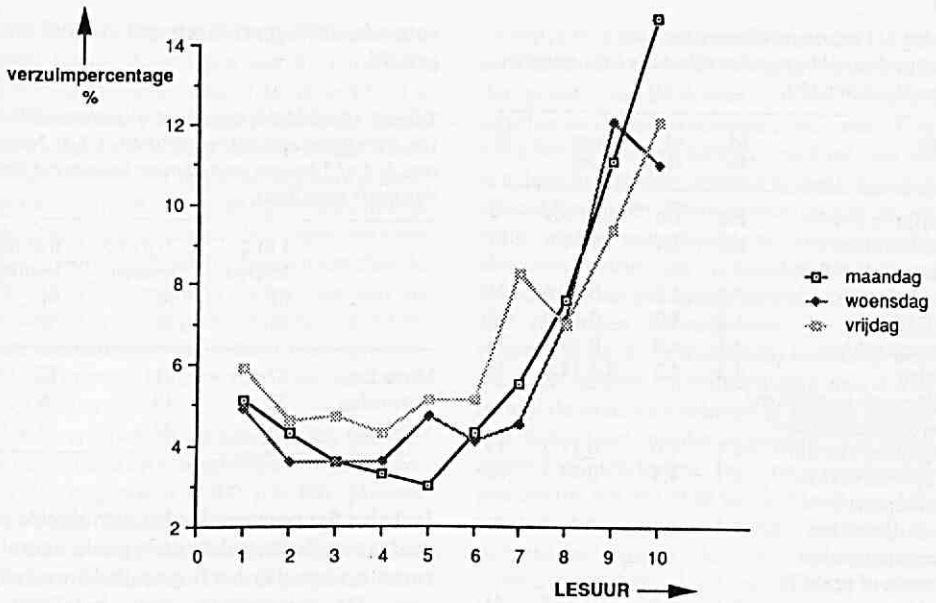
In Figuur 3 zijn de verschillen tussen de gemiddelde verzuimpercentages per lesuur op de drie meetdagen in beeld gebracht.

De lesuren zijn op de horizontale as en de gemiddelde verzuimpercentages per lesuur per dag op de verticale as afgebeeld. De 'vrijdag'-lijn ligt vanaf het eerste tot en met het zevende lesuur boven de lijnen voor maandag en woensdag. De gemiddelde verzuimpercentages voor de lesuren op maandag en woensdag lopen uiteen. Dit geldt vooral voor het tweede, vijfde en zevende lesuur. Op maandag wordt er gemiddeld per school op het tweede, zevende, achtste en tiende lesuur meer gespijbeld dan op woensdag op dezelfde lesuren. Het omgekeerde geldt voor het vijfde lesuur. Op de andere lesuren zijn de verschillen gering.

Een mogelijke verklaring voor het verschil met betrekking tot het tweede lesuur zou kunnen zijn dat op maandag (na het weekend) sprake is van meer 'langslapers' dan op woensdag. Ook voor het vijfde lesuur zijn verklaringen voor het verschil in gemiddelde verzuimomvang tussen maandag en woensdag denkbaar. Het zou bijvoorbeeld zo kunnen zijn dat op woensdag meerdere leerlingen 'eerder middagpauze nemen'. De standaarddeviatie voor de gemiddelde verzuimomvang



Figuur 2 *Gemiddelde verzuimpercentages per lesuur (2<N<35)*



Figuur 3 *Gemiddelde verzuimpercentages op de verschillende lesuren op maandag, woensdag en vrijdag (2<N<35)*

van woensdag het vijfde lesuur (7,0) is veel hoger dan die van de meeste andere lesuren op de drie dagen (veelal rond de 4,0). Het verschil tussen beide dagen zou daarom door een relatief gering aantal scholen kunnen worden veroorzaakt.

Verzuimomvang per vak

Er bestaan veel speculaties over de mate waarin er bij verschillende vakken gespijbeld zou worden. Op basis van de in het onderhavige onderzoek verzamelde verzuimgegevens kan de houdbaarheid van dergelijke veronderstellingen worden vastgelegd.

De verzuimpercentages zijn voor elke school per vak berekend. Daartoe zijn verschillende groepen van vakken geconstrueerd. Deze zijn te vinden in Tabel 4. De exacte vakken (wis-, natuur- en scheikunde, biologie) en de talen (waaronder Nederlands en de moderne vreemde talen) zijn samengenomen, evenals de zaakvakken (zoals aardrijkskunde en geschiedenis) en de expressievakken (onder andere handenarbeid, tekenen en lichamelijke opvoeding). Met beroepsgerichte theorie- of praktijkvakken worden beroepsgerichte vakken bedoeld, waarvan het onduidelijk is of het een theorie- dan wel een praktijkvak is.

Bepaalde vakken worden op zowel lbo- als avo-scholen gegeven, andere worden slechts op één van beide schooltypen gegeven. Dit is de belangrijkste reden voor de verschillen in het aantal scholen (N) per vakkencombinatie.

Uit Tabel 4 blijkt dat de gemiddelde omvang van het verzuim bij de economische vakken en bij de beroepsgerichte (met name de theorie-) vakken meer dan 5% bedraagt. Bij de economische vakken valt de relatief hoge spreiding onder de 22 scholen op. Een relatief geringe verzuimomvang is geconstateerd bij de vakken Latijn en Grieks (2.2%), bij het vak 'onderwijs in de eigen taal en cultuur' (oetc; 3.8%), bij de zaakvakken (3.3%) en bij maatschappijleer en godsdienst (3.7%). Bij de klassieke talen moet worden opgemerkt dat deze slechts op 4 van de 35 onderzochte scholen worden onderwezen. Het gemiddelde verzuimpercentage voor de eerste 4 in de Tabel genoemde vakken is lager dan de percentages voor de exacte vakken en de talen. In de onderwijspraktijk wordt vaak het tegendeel verondersteld.

De gemiddelde verzuimomvang van de andere vakken ligt rond de gemiddelde verzuimomvang van de 35 scholen die in Tabel 1 is weergegeven (4.4%).

De hoogste gemiddelde verzuimpercentages zijn dus geconstateerd voor typische lbo-vakken, maar ook voor vakken die vooral op

Tabel 4 *Verzuimpercentage per vak. Gemiddelde (M), standaarddeviatie (sd), range en aantal scholen (N)*

Vak	M %	sd %	range %-%	N
Latijn en Grieks	2.2	1.6	0.3- 3.9	4
Zaakvakken	3.3	3.4	0-14.6	33
Maatschappijleer/ Godsdienst	3.7	4.3	0-15.6	24
OETC	3.8	3.7	0-14.0	35
Exacte vakken	4.1	4.0	0-16.9	35
Talen	4.2	4.2	0.2-15.0	34
Algemene technieken/ beroepenoriëntatie/ algemeen avo-uur	4.4	4.0	0-11.2	19
Expressievakken	4.7	4.3	0-17.3	35
Beroepsgerichte praktijkvakken	5.5	4.6	0-15.0	31
Beroepsgerichte theorie of praktijk-vakken	6.1	5.8	0-23.5	21
Economische vakken	7.0	7.6	0-26.9	22
Beroepsgerichte theorievakken	7.8	6.7	0-27.2	21

avo-scholen worden gegeven: de economische vakken.

De toch relatief lage gemiddelde verzuimpercentages voor de exacte vakken kunnen wellicht als volgt verklaard worden. Leerlingen vinden deze vakken doorgaans moeilijk en spijbelen daarom bij voorkeur niet bij deze vakken, maar bij andere vakken die ze minder moeilijk of belangrijk vinden.

Het relatief hoge gemiddelde verzuimpercentage voor de economische vakken (voor een deel exact) is in dit licht bezien extra opvallend.

5.3 *Aanvullende data over de verzuimomvang*

De hierboven gepresenteerde verzuimpercentages geven geen inzicht in het aantal uren dat individuele leerlingen spijbelden. Om een beeld te krijgen van het aantal uren dat leerlingen spijbelden is het gemiddelde aantal gespijbelde uren per spijbelaar voor elke meetdag berekend. Omdat de namen van de spijbelaars via de absentieformulieren bekend waren is per school per spijbelaar het aantal uren dat hij of zij op de drie meetdagen ongeoorloofd afwezig was, berekend. Aan de hand van deze gegevens is ook het absolute aantal spijbelaars voor de drie meetdagen per school vastgesteld.

Tabel 5 *Gemiddeld percentage spijbelaars (M) van alle spijbelaars dat respectievelijk 1 of 2 lesuren, 3, 4 of 5 lesuren en 6 of meer lesuren per meetdag heeft gespijbeld.*

	1 of 2 lesuren M %	3, 4 of 5 lesuren M %	6 of meer lesuren M %
Maandag	67	21	12
Woensdag	72	19	9
Vrijdag	64	33	3

In Tabel 5 is per meetdag het gemiddelde percentage spijbelaars dat een bepaald aantal lesuren op een dag heeft gespijbeld weergegeven. De percentages zijn berekend in verhouding tot het totale aantal verschillende leerlingen dat per meetdag op een school heeft gespijbeld. Op maandag, woensdag en vrijdag wordt gemiddeld op de scholen door ongeveer *tweederde* van alle spijbelaars 1 of 2 lesuren verzuimd.

Een relatief kleine groep spijbelaars is gedurende 6 of meer lesuren waarop hij of zij aanwezig zou moeten zijn ongeoorloofd afwezig (ongeveer 12% op maandag, 9% op woensdag en 3% op vrijdag). Opvallend is dat deze groep leerlingen op maandag het grootst is en afneemt naarmate de week vordert. In Tabel 5 valt het verschil in het gemiddelde percentage spijbelaars dat 3, 4 of 5 lesuren spijbelde tussen enerzijds maandag en woensdag (21% en 19%) en anderzijds vrijdag (33%) op. Eerder is in de beschrijving van de omvang van het schoolverzuim geconstateerd dat op vrijdag gedurende de eerste 7 lesuren meer wordt gespijbeld dan op de twee andere meetdagen. Tesamen met de hier genoemde resultaten kan worden verondersteld dat dit verschil vooral wordt bepaald door de toename van het percentage spijbelaars dat 3,4 of 5 lesuren spijbelde.

Een tweede maat die meer inzicht kan geven in de samenstelling van de verzuimomvang is de spijbelaarsratio: de verhouding tussen het totaal aantal verschillende spijbelaars dat op de drie meetdagen is geregistreerd en het totale aantal leerlingen dat bij de verzuimmeting was betrokken. Beide aantallen zijn

voor elke school bepaald. De spijbelaarsratio varieert tussen de scholen van 0.01 tot 0.51 met een gemiddelde van 0.16 (N = 34). De standaarddeviatie bedraagt 0.11. Deze resultaten wijzen op grote verschillen tussen scholen voor wat betreft het absolute aantal leerlingen dat spijbelt. Waar de ene school in drie dagen 2 spijbelaars registreert, registreert een andere school op dezelfde dagen meer dan de helft van alle leerlingen gedurende één of meerdere lesuren als spijbelaar. De verschillen tussen scholen voor wat betreft hoe men spijbelen definieert, kunnen hier uiteraard ook een rol spelen.

Het correlationele verband tussen het verzuimpercentage op schoolniveau en de spijbelaarsratio is groot ($r = .81$; $p < .00$). Hieruit kan worden geconcludeerd dat op scholen waar relatief (ten opzichte van de andere scholen in dit onderzoek) veel wordt gespijbeld een relatief grote groep leerlingen voor deze verzuimomvang verantwoordelijk is.

6 *Conclusies*

De verzamelde gegevens geven een nauwkeurig beeld van de omvang van het schoolverzuim op 36 scholen in vier grote Nederlandse steden. Door het selectieve karakter van de onderzoeksgroep kunnen de resultaten van het onderzoek echter niet worden gegeneraliseerd naar het gehele Nederlandse voortgezet onderwijs. Het betreft, vanwege de project gebondenheid van de scholen, geen a-selecte steekproef. In de onderzoeksgroep zijn kleine scholen en lbo/lbo-scholengemeenschappen oververtegenwoordigd; bovendien zijn veel scholen gevestigd in de drie grootste gemeenten van Nederland. De resultaten van dit onderzoek kunnen misschien als een indicatie voor de bovengrens van de verzuimomvang in het Nederlandse voortgezet onderwijs worden opgevat omdat verwacht mag worden dat de verzuimomvang in de rest van Nederland niet hoger zal zijn dan in deze steden.

Welke indicaties leverde de verzuimmeting op?

De op schoolniveau berekende verzuimpercentages wezen op redelijk grote verschillen tussen de scholen. Het gemiddelde verzuimpercentage van alle in het onderzoek betrokken scholen bedraagt 4.4%. Opvallend is dat het gemiddelde percentage geoorloofde absentie hoger uitvalt (4.7%) en dat de verschillen tussen de scholen wat dit betreft minder groot zijn. Blijkbaar zijn op de meeste scholen leerlingen minstens even vaak of vaker afwezig met een geldige reden (al dan niet een door hun ouders gedekte smoes), dan zonder geldige reden. Een combinatie van de omvang van het geoorloofde en ongeoorloofde absenteïsme leert dat op de scholen uit de onderzoeksgroep per lesuur gemiddeld 9 van de 100 leerlingen afwezig zijn. De lbo/lbo-scholen en de lbo/avo-scholen hebben gemiddeld een hogere verzuimomvang dan 4.4%, terwijl de avo/vwo-scholen te maken hebben met beduidend minder ongeoorloofde afwezigheid van leerlingen. In bijna alle schooltypen neemt het verzuim per leerjaar toe. Uitzonderingen op deze regel vormen de overgangen tussen de schooltypeleerjaren havo-2 naar havo-3 en vwo-3 naar vwo-4. Voor alle schooltypen geldt dat in het voorexamenleerjaar in vergelijking met andere leerjaren in dat schooltype het meest gespijbeld wordt. Dit kan deels samenhangen met de leeftijd van de leerlingen. Zodra leerlingen niet meer leerplichtig zijn zouden scholen toleranter kunnen worden en leerlingen zich meer verzuim permitteren.

De verschillen tussen enerzijds ibo- en lbo-leerjaren en anderzijds avo- en vwo-leerjaren komen naar voren bij verzuimpercentages die berekend zijn per schooltypeleerjaar. Voor het tweede en derde leerjaar is een vergelijking tussen beide groepen mogelijk. Hieruit blijkt dat in de ibo- en lbo-groepen in het tweede en derde leerjaar veel meer wordt gespijbeld dan in de beide andere schooltypen. Dit zou samen kunnen hangen met de sociaal-economische achtergrond (onder andere huisvesting en gezinsinkomen) van de leerlingen, die van de ibo/lbo-leerlingen veelal zwakker wordt geacht. Uit andere analyses is gebleken dat het aantal buitenlandse (met name Turkse en Marokkaanse) leerlingen, dat groter is op ibo/lbo-scholen dan op avo/vwo-scholen, positief samenhangt met de verzuimomvang (Bos et al., 1989). Dit hoeft nog niet te betekenen dat vooral de allochtone leerlingen spijbelen (vgl. De Vries & Peetsma, 1987). Nader onderzoek zou kunnen uitwijzen dat op scholen waar relatief veel buitenlandse leerlingen staan ingeschreven het juist de Nederlandse leerlingen zijn die spijbelen.

De gemiddelde verzuimomvang, berekend

per dag, verschilt voor maandag en woensdag niet veel van elkaar. Op vrijdag wordt er gemiddeld meer gespijbeld dan op maandag en woensdag. Uit de gegevens per lesuur blijkt dat met name tijdens de eerste zeven lesuren op vrijdag meer gespijbeld wordt dan op dezelfde lesuren op maandag en woensdag. Leerlingen spijbelen op alle meetdagen op het eerste lesuur aanzienlijk meer dan op de drie daaropvolgende uren. Vanaf het vijfde lesuur neemt het spijbelen weer toe, behalve op maandag. Na het zesde uur neemt de verzuimomvang aanmerkelijk toe.

Ten slotte is de verzuimomvang per vak vastgesteld. Hierbij vallen relatief (ten opzichte van het scholengemiddelde) lage en relatief hoge verzuimpercentages op. Relatief lage verzuimpercentages zijn geconstateerd bij zaakvakken, de klassieke talen, het vak onderwijs in eigen taal en cultuur en bij maatschappijleer en godsdienst. Vooral het verzuimpercentage voor maatschappijleer en godsdienst is tegenovergesteld aan wat doorgaans wordt verondersteld. Ook bij de expressievakken is het verzuim lager dan vaak aangenomen wordt (4.7%). Relatief hoge verzuimpercentages werden vastgesteld voor de beroepsgerichte vakken en de economische vakken. De beroepsgerichte vakken worden voornamelijk gegeven op de lbo-schooltypen. Voor de relatief hoge verzuimomvang bij de economische vakken, die vooral op avo/vwo-scholen worden onderwezen, zou een gering aantal scholen (outliers) verantwoordelijk kunnen zijn aangezien de spreiding hier hoog bleek te zijn.

In dit onderzoek is de samenhang tussen elk van de onafhankelijke variabelen (bijvoorbeeld lesuur, schooldag en vak) *apart* met de afhankelijke variabelen (de verzuimomvang) onderzocht. Een probleem bij de interpretatie van deze samenhangen is dat de onafhankelijke variabelen onderling ook samen kunnen hangen. Stel bijvoorbeeld dat de variabele lesuur sterk samenhangt met spijbelen, maar de variabele lesuur eveneens sterk samenhangt met de onafhankelijke variabele vak. De vraag rijst dan in hoeverre er nog een samenhang tussen lesuur en spijbelen blijkt te zijn, als de variabele vak constant gehouden wordt. Er is dus behoefte aan een model dat de relatieve bijdrage van de diverse variabelen aan het verzuim weergeeft. Het is daarom de bedoeling om een secundaire analyse in de vorm van een log-lineaire analyse op de beschikbare data uit te voeren.

Uit de berekeningen met betrekking tot de gemiddelde omvang van het verzuim *per spijbelaar* blijkt onder meer dat ongeveer tweederde van alle spijbelaars op zowel maandag, woensdag als vrijdag één of meer dan twee lesuren ongeoorloofd afwezig is. Op vrijdag is het percentage spijbelaars dat 3,4 of 5 lesuren spijbelt hoger (± 20%) dan dat op maandag en woensdag (± 33%). Het percentage *leerlingen* dat op de drie meetdagen één of meer lesuren ongeoorloofd afwezig is geweest (de spijbelaarsratio) varieert van 1% tot 51%, met een gemiddelde van 16%.

Deze resultaten overziend kan worden geconcludeerd dat het schoolverzuim in vier grote steden een grote groep leerlingen betreft. Door een meerderheid van de spijbelaars wordt per dag niet meer dan twee lesuren verzuimd, maar dit aantal is toch aanzienlijk. Het geconstateerde verzuim leidt in combinatie met de omvang van de netto lesuitval (9.5%, volgens een onderzoek dat is uitgevoerd door het Ministerie van Onderwijs en Wetenschappen, 1989) tot veel niet genoten onderwijs. De invloed daarvan op de scholing en vorming van leerlingen verdient serieuze aandacht. Het lijkt daarom zinnig om in de school te blijven streven naar terugdringing dan wel voorkoming van verzuim.

Door de landelijke overheid wordt veel aandacht geschonken aan schoolverzuim in de vorm van studies naar de omvang ervan en het stimuleren van initiatieven om het verzuim terug te dringen. De kennis ten aanzien van het verzuimprobleem is echter nog beperkt, mede omdat het een fenomeen is dat moeilijk meetbaar is. Om het inzicht in de verzuimproblematiek te vergroten zal naar onze mening moeten worden gestreefd naar verder grootschalig en langdurig onderzoek op dit terrein. Vermoedens dat het verzuim de laatste jaren zou zijn afgenomen en nieuw verzuimonderzoek daarom niet nodig zou zijn (zie onder andere De Vries & Peetsma, 1987), hebben een zwakke empirische basis.

Dergelijk onderzoek zou ook gericht moeten zijn op het vaststellen van verzuimpatronen van spijbelaars (incidenteel, kortstondig met patroon, periodiek, permanent verzuim; zie het inspectierapport uit 1989). Bovendien zou dergelijk onderzoek moeten uitwijzen hoe de verzuimomvang varieert over diverse

momenten in het schooljaar en hoe het beeld eruit ziet voor verschillende delen van Nederland (inclusief de vergelijking tussen stad en platteland). Naast de vergroting van de kennis over de verzuimomvang en patronen daarin door middel van langdurige exacte metingen is verder onderzoek naar de achtergronden van het verzuim noodzakelijk voor het kunnen nemen van meer doelgerichte verzuimreducerende maatregelen.

Literatuur

Babeliowsky, M., *Voortijdig schoolverlaten in Amsterdam*. Een onderzoek naar het voortijdig schoolverlaten en schoolverzuim in het Amsterdamse voortgezet en middelbaar beroepsonderwijs. Amsterdam: SCO, 1986.

Baum, Th., Surveys of absenteeism: a question of timing. *Educational Research*, 1978, vol. 20, no. 3, 226-230.

Billington, B. J., Truants: some personality characteristics. *Durham and Newcastle Research Review*, 1979, vol. 9, no. 43, 1-6.

Bos, K. Tj., A. M. Ruijters, A. J. Visscher, *Some school effectiveness indicators and their relation with school characteristics*. Paper gepresenteerd op de Third Joint Meeting OTG-Onderwijsleerprocessen, Enschede, 25-4-1989.

Bos, K. Tj., B. J. van Kesteren, W. G. R. Stoel & C. J. A. J. Vermeulen, *Schoolverzuim: maatregelen die scholen (kunnen) nemen*. Amsterdam: Swets & Zeitlinger, 1990.

Dorsteen, T., *Spijbelen*. Uitslag vragenlijst grote 'Je ziet maar'-spijbelenquête. Hilversum: VARA, 1985.

Galloway, D., *School and persistent absentees*. Oxford: Pergamon Press, 1985.

Geilman, G. J., *Verslag van een onderzoek naar de aanwezigheid op school*. Rotterdam: Bureau leerplichtzaken, 1984.

Ginjaar-Maas, N., *Beleidsnotitie schoolverzuim*. 's-Gravenhage: Ministerie van Onderwijs en Wetenschappen, 1986.

Grimshaw, R. & J. Pratt, Counting the absent scholars: some implications for managerial practice arising from a survey of absenteeism in a city's secondary school. *School Organization*, 1986, vol. 6, no. 1, 155-173.

Hoen, R. H. A. 't, P. A. M. Bosch & W. Verhagen, *Met het oog op stiemen*. Een onderzoek naar schoolverzuim en voortijdig schoolverlaten in Arnhem. Arnhem: Gemeente Arnhem, 1986.

Inspectie van het Onderwijs, *Bestrijding van het schoolverzuim in het voortgezet onderwijs; resultaten van een inspectie-onderzoek*. 's-Gravenhage: Ministerie van Onderwijs en Wetenschappen, 1989.

Jaspers, J., H. Schut, J. de Keyser, J. W. Gunnink & F. Hüsken, *Van school gaan*. Leeuwarden: Hoger Onderwijs Friesland, 1984.

Karweit, N., *Rainy days and Mondays*. An analysis of factors related to absence from school. Baltimore (Maryland): John Hopkins University, 1973.

Lesman, T., *Schoolverzuim*. Een onderzoek naar aard, omvang en schoolfaktoren in het individueel beroepsonderwijs. Amsterdam: ABC, 1980.

Ministerie van Onderwijs en Wetenschappen, *Lesuitval, een verkennend onderzoek naar aard en omvang van lesuitval in het voortgezet onderwijs*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1989.

Reid, K., *Truancy and School Absenteeism*. Londen: Hodden and Stoughton, 1985.

Stoel, W. G. R., *Schoolkenmerken en het gedrag van leerlingen en docenten in het voortgezet onderwijs*. Groningen: RION, 1986.

Vlug, I., *Schoolverzuim van Turkse en Marokkaanse meisjes*. Rotterdam: Erasmus Universiteit, vakgroep Onderwijssociologie en -beleid, 1985.

Vries, G. C. de & M. W. de Jong, *Schoolverzuim en schooluitval*. Achtergronden, feiten en verklaringen in vogelvlucht. Hoevelaken: LPC, 1986.

Vries, G. C. de & Th. Peetsma, *Voortijdig schoolverlaten en spijbelen in het voortgezet onderwijs*. Resultaten van een onderzoek betreffende het schooljaar 1984-1985. Amsterdam: SCO, 1987.

Vries, G. C. de, G. Monsma & E. Mellink, *Anatomie van het schoolverzuim*. Inleiding conferentie Schoolverzuim en onderwijsbeleid van de landelijke contactgroep schoolverzuim. Zeist, 15 februari 1989.

Curricula vitae

K. Tj. Bos is ontwikkelingspsycholoog en als onderwijskundig onderzoeker werkzaam binnen het Onderzoek Centrum Toegepaste Onderwijskunde (OCTO) van de Universiteit Twente. Hij doet momenteel onderzoek naar schoolverzuim in het voortgezet onderwijs en naar de effecten van nascholingscursussen op het gebied van onderwijskundig management in het basisonderwijs.

A. M. Ruijters is onderwijskundige en als onderzoeker werkzaam geweest binnen het Onderzoek Centrum Toegepaste Onderwijskunde (OCTO) van de Universiteit Twente. Zij heeft onderzoek gedaan naar het gebruik en de effecten van een computerondersteund absentieregistratiesysteem en naar breedschalig computergebruik in het Proefstation Oost Nederland.

A. J. Visscher is onderwijskundige en als universitair docent werkzaam bij de vakgroep Onderwijsorganisatie en -management van de Faculteit Toegepaste Onderwijskunde van de Universiteit Twente. Hij doet onderzoek naar en publiceert over de mogelijkheden van computerondersteunde schoolinformatiesystemen.

Adres: Faculteit Toegepaste Onderwijskunde, Universiteit Twente, Postbus 217, 7500 AE Enschede

Manuscript aanvaard 16-2-'90

Summary

Bos, K. Tj., A. M. Ruijters & A. J. Visscher. 'Absenteeism in secondary education'. *Pedagogische Studiën*, 1990, *67*, 341-354.

In spite of the large attention given to the truancy problem empirical data concerning this phenomenon are limited. In this article we report on a study of the extent of the lawful absenteeism and truancy of pupils at thirty-six schools in four big Dutch cities. Truancy was determined at eight levels (amongst others for each school, grade, subject) during three schooldays and 8990 lesson periods. The results show that the lawful absenteeism (4.7%) is on average somewhat higher than the truancy magnitude (4.4%). Between schools considerable differences with regard to the extent of truancy exist. At schools with a lower vocational education department on average more pupils play truant. The same goes for most of the higher grades, friday, the first and the seventh up to and including the tenth lesson period, the professionaly orientated and economic subjects. For subjects that often are regarded as 'truancy-subjects' relatively low truancy percentages were measured. Most of the truants play truant no more than two lesson periods at one day. An average of sixteen percent of the pupils of a school plays truant, although this differs strongly between schools.

Systematisering van de analyse in de gefundeerde theorie-benadering; een illustratie

R. RICHARDSON, E. JONGELING en F. WESTER
Vakgroep Algemene Pedagogiek
Katholieke Universiteit Nijmegen

Samenvatting

Ondanks de methodologische richtlijnen voor kwalitatief onderzoek, ontbreken voor onderzoekers vaak nog steeds scherpe onderzoeksvoorschriften. Dit geeft vaak aanleiding tot valkuilen en problemen bij de toepassing van kwalitatief onderzoek. De auteurs van dit artikel hebben getracht dit probleem zoveel mogelijk op te lossen met betrekking tot een van de meest toegepaste strategieën voor kwalitatief onderzoek: de gefundeerde theorie-benadering. Zij beschrijven een aantal analyse-stappen als richtlijn voor systematische analyse van het materiaal met het oog op begripsontwikkeling. Belangrijk hulpmiddel tijdens de uitvoering van de stappen is het computerprogramma KWALITAN. Een en ander wordt geïllustreerd met gegevens uit het onderzoek van Jongeling: "Perspectieven op onderwijs van leerlingen uit het Voortgezet Onderwijs".

1 *Inleiding*

De laatste tien jaren kan men een groeiende interesse in de methoden voor kwalitatief onderzoek constateren. Dit heeft onder meer geleid tot een veel grotere aandacht voor de analyse van kwalitatief materiaal (Glaser, 1978; Maso, 1987; Miles & Huberman, 1984; Strauss, 1987; Wester, 1984, 1987). We denken daarbij aan procedures als 'open coderen', 'theoretical sampling', 'constant vergelijken', 'matrix analyses'.

Ondanks deze concretiseringen van methodologische richtlijnen blijft de analyse van kwalitatief materiaal in praktische zin voor veel onderzoekers nog erg onduidelijk. De vraag blijft via welke concrete onderzoekshandelingen men dergelijke procedures gestalte kan geven.

In deze bijdrage stellen wij een stapsgewijze procedure van kwalitatieve analyse voor, die is gerelateerd aan een van de meest toegepaste strategieën voor kwalitatief onderzoek: de Gefundeerde Theorie-benadering (verder g.t.-benadering) van Glaser en Strauss. In hun boek "Discovery of Grounded Theory" (1967) leggen zij de nadruk op de stapsgewijze ontwikkeling van theorie door middel van een grondige kennismaking met de werkelijkheid. Aan het begin van het onderzoek wordt gebruik gemaakt van globale inzichten, waaraan de attenderende en richtinggevende begrippen ('sensitizing concepts') worden ontleend, die door voortdurende confrontatie met de empirie substantief worden gemaakt en verscherpt in de definiëring. Belangrijke principes zijn 'theoretical sampling' en 'constante vergelijking'. Wanneer men op het spoor is gekomen van belangrijke elementen van de zich ontwikkelende theorie, gaat men op zoek naar nieuwe gevallen, die de eerdere bevindingen kunnen verscherpen, bevestigen of corrigeren. De analyse bestaat dan ook uit een constante vergelijking van voorbeelden. Dit vergelijkingsproces gaat zolang door tot geen nieuwe inzichten meer worden verkregen.

Tot zover deze korte weergave van de onderzoeksopvattingen van Glaser en Strauss. Naast enthousiasme, zijn er kritische reacties gekomen op 'Discovery'. De voornaamste problemen die men voorziet zijn het gevaar voor inductivisme (een misverstand opgeroepen door het polemisch karakter van Discovery) en de vaagheid van de onderzoeksvoorschriften (Wester, 1984).

Met name het laatste punt kan aanleiding geven tot valkuilen en problemen bij de toepassing van de g.t.-benadering. Het gevaar bestaat dat het ontbreken van een scherpe probleemstelling, van een vaststaand theoretisch kader, of van een systematische waarnemingsprocedure betekent het ontbreken van elk houvast. Na verloop van tijd ontdekt de

onderzoeker dan dat de samenhang ontbreekt in de onderzoekshandelingen, dat men eeuwig kan doorgaan met het ontdekken van nieuwe categorieën of het stellen van nieuwe vragen, dat de analyse ongericht is of de relevantie van de gegevens vaag blijft.

Dit soort ervaringen is aanleiding geweest tot het verder uitwerken van deze benadering (Wester, 1984, 1987), waarbij de nadruk kwam te liggen op de volgende punten:

1. Bij elk analyse-proces gaat het om een wisselwerking tussen begrippen en gegevens. In de g.t.-benadering is dat heel sterk het geval, omdat de begrippen in de loop van het onderzoek uitgewerkt worden. Een centraal principe is nu dat met de cognitieve problemen die dit oplevert wordt omgegaan *door ze centraal te stellen!* De reflectie op onderzoeksvragen, dataverzameling, analyse en begrippenkader krijgt daartoe dagelijks aandacht en wordt vastgelegd in memo's (korte notities). De onderzoekshandelingen reflectie, dataverzameling en analyse vormen een cyclus, die meermalen wordt doorlopen, gestuurd door het reflectieproces.
2. Het proces van het ontwikkelen van de theorie kan uiteengelegd worden in vier fases. In elke fase wordt de cyclus enkele malen doorlopen om het centrale doel van de fase te bereiken. In de exploratiefase wordt het terrein verkend, zoveel mogelijk ideeën opgedaan, voorlopige begrippen (aspecten, eigenschappen, kenmerken, variabelen, categorieën, indicatoren) geformuleerd en als trefwoorden vastgelegd aan het materiaal. In de specificatie fase worden deze begrippen uitgewerkt en zo mogelijk onder de meest centrale begrippen geordend. In de reductiefase wordt gericht gezocht naar een centrale optiek (kernbegrip of basisproces) van waaruit het geheel kan worden geordend. In de integratiefase worden de relaties tussen de centrale begrippen in de ontwikkelde optiek systematisch onderzocht.

In deze bijdrage willen wij concrete stappen voor de analyse van het materiaal in de eerste twee fasen presenteren en illustreren.

Bij deze concretisering van het analyseproces zijn twee zaken van belang geweest. Ten eerste de ervaringen van diverse onderzoekers met de computer als hulpmiddel bij de analyse van het kwalitatieve materiaal. De toepassing van de computer dwingt de onderzoeker ertoe de verschillende onderzoekshandelingen te expliciteren en te systematiseren. De computer analyseert niet, maar is een belangrijk hulpmiddel bij het organiseren van de papierwinkel: het opslaan, bewerken, ordenen en selecteren van het vaak omvangrijke materiaal (zie voor de rol van de computer Peters & Wester, 1988).

Ook van belang was de onderzoekspraktijk van diverse kwalitatieve, op theorieontwikkeling gerichte onderzoeken, in Nederland. De concretisering van de analyseprocedures tot analyse-stappen heeft in nauw contact met die onderzoekspraktijk plaatsgevonden. Een belangrijk onderzoek voor de ontwikkeling van de analyse-stappen is het onderzoek 'Perspectieven op onderwijs van leerlingen uit het voortgezet onderwijs' dat door Jongeling wordt uitgevoerd aan de Rijksuniversiteit Utrecht. De ontwikkelde analyse-stappen zullen wij in de volgende paragrafen illustreren met gegevens uit het bovengenoemde onderzoek van Jongeling. In de slotparagraaf zullen enkele discussiepunten met betrekking tot de analyse van materiaal volgens de principes van de g.t.-benadering aan de orde worden gesteld.

2 *Stappen in de exploratiefase*

Voordat de analyse van het materiaal begint heeft al het een en ander plaats gevonden. Ten eerste is het onderzoek gestart met een al of niet ver uitgewerkte probleemstelling en bijbehorende onderzoeksopzet. Ten tweede is er een begin gemaakt met de verzameling van materiaal. In feite betekent dit dat de onderzoeker een aantal keuzen heeft gedaan waarin onderdelen van het analytisch kader (begrippen, variabelen, onderzoeksthema's, gesprekstopics) een rol hebben gespeeld. De eerste stap in de exploratiefase bereidt de analyse voor door te inventariseren om welke begrippen[1] het gaat en welke vragen in verband hiermee aan het materiaal worden gesteld. De tweede stap in de exploratiefase is gericht op het opdoen van zoveel mogelijk ideeën omtrent wat er speelt op het empirisch veld.

2.1 *Stap 1: Voorbereiding van de data-analyse*

In paragraaf 1 hebben we gezien dat een van de kenmerken van de g.t.-benadering is dat de onderzoeker de analyse start met open en algemene begrippen die worden geordend in een conceptueel raamwerk. Deze begrippen alléén zijn ontoereikend om een koppeling te kunnen maken met het onderzoeksmateriaal. Om de kloof tussen begrippen en data te kunnen overbruggen worden concrete onderzoeksvragen geformuleerd. Met behulp van het conceptueel raamwerk wordt vervolgens bepaald bij welke eenheden data worden verzameld en welke instrumenten daarbij worden gebruikt.

Ter voorbereiding van de analyse worden dan de volgende handelingen verricht, die in deze paragraaf achtereenvolgens zullen worden beschreven aan de hand van het onderzoek naar perspectieven op onderwijs van leerlingen uit het voortgezet onderwijs:

a. Specificeren van de achtergronden van het onderzoek;
b. Opbouw van het conceptueel raamwerk;
c. Specificeren van de onderzoeksvragen;
d. Keuze van de eenheden;
e. Instrumentatie en eerste dataverzameling;
f. Overzichtmemo's (sensitizing concepts en analyse-vragen).

Als laatste handeling wordt hierboven genoemd het schrijven van overzichtsmemo's. Memo's zijn een belangrijk hulpmiddel in de kwalitatieve analyse. Zij bevatten een weergave van reflecties van de onderzoeker naar aanleiding van de waarneming en de analyse, en sturen daarmee het onderzoek en dus ook de analyse. Het cyclisch karakter van het onderzoeksproces en de centrale plaats van reflectie daarin zorgen ervoor dat de onderzoeker voortdurend memo's zal produceren.

Memo's verschillen wat betreft de functie die zij vervullen in de kwalitatieve analyse (vgl. Peters, Wester & Richardson, 1989). Ten eerste zijn er *theoretische* memo's waarin beslissingen met betrekking tot de centrale elementen van de theorie worden opgenomen. Ten tweede zijn er *methodische* memo's waarin keuzen van de onderzoeker met betrekking tot de voortgang van het onderzoek worden vastgelegd. Deze beide soorten memo's hebben direct te maken met het reflecteren op de resultaten van waarneming en analyse en het gebruik van inzichten ontleend aan andere bronnen zoals literatuur of gesprekken met collega's. Een andere soort memo's zijn *begrippenkaarten* om de betekenis vast te leggen van al de 'begrippen' die in de loop van de analyse worden gehanteerd. Daarnaast zijn er *profielkaarten* die bedoeld zijn om de onderzochte eenheden te typeren op basis van de meest centrale begrippen. In een latere fase van de analyse, met name tijdens het analyseren van overzichtskaarten spelen deze profielkaarten een belangrijke rol. Een laatste en naar functie afwijkende soort memo's zijn de *overzichts- of recapitulatiememo's*. Dit zijn theoretische of methodische memo's, waarin enkele voorgaande memo's over hetzelfde thema worden samengevat en een conclusie wordt getrokken. In feite worden overzichtsmemo's geschreven naar aanleiding van een analyse van die voorafgaande memo's.

In deze fase van het onderzoek, de analysevoorbereidende fase, worden ook overzichtsmemo's geschreven. In die overzichtsmemo's worden de belangrijkste aantekeningen met betrekking tot de probleemstelling van het onderzoek samengebracht. Daarin zijn ook de belangrijkste aantekeningen verwerkt met betrekking tot thema's of begrippen die men eventueel naar aanleiding van literatuurstudie heeft opgedaan. Op deze wijze zijn overzichtsmemo's mede gebaseerd op literatuurstudie, dan wel aanleiding voor nadere literatuurstudie om de probleemstelling scherper te kunnen formuleren. Overzichtsmemo's komen tot stand door allereerst achtergronden en conceptueel raamwerk van het onderzoek te formuleren. Hierdoor vormt de onderzoeker zich een beeld van het onderzoeksveld en brengt deze de belangrijkste begrippen in kaart. De achtergronden van een onderzoek worden gevormd door antwoord te vinden op vragen als:

1. Waarom is een bepaald onderzoek interessant?
2. Welke lacunes zijn er in het onderzoeksveld te ontdekken?
3. Op welke wijze denkt men die lacunes in te vullen?

De achtergronden en het conceptueel raamwerk zijn als volgt beschreven in de onderzoeksaanvrage van het onderzoek naar onderwijsperspectieven:

De aanleiding voor het onderzoek vormen de vele problemen verbonden aan de participatie van leerlingen aan hun onderwijs. Verschijnselen als het voortijdig de school verlaten, regelmatig van school spijbelen en lesondergravend gedrag hebben de afgelopen decennia een zorgwekkende omvang aangenomen. Vanuit een sociologische optiek worden deze problemen in verband gebracht met bredere maatschappelijke ontwikkelingen (o.a. Bernards, 1986; Du Bois-Reymond & Veld, 1986; Matthijssen, 1986; Van Gelder & Tillekens, 1986). Daarbij wijzen de auteurs op de problemen van aansluiting van het onderwijs op de arbeidsmarkt met de toegenomen (jeugd)werkloosheid en de verzwakking van de relaties tussen schoolse kennis en beroepsfracties. Ook wordt de aandacht gevestigd op de kloof tussen enerzijds de toenemende zelfstandigheid van jongeren in buitenschoolse leefverbanden en anderzijds de vaak binnen school verlangde passief consumerende houding. Ten slotte wordt door auteurs gewezen op de toenemende concurrentie die het onderwijs ondervindt vanuit de vrije-tijdscultuur. Deze maatschappelijke ontwikkelingen lijken schoolgaan voor jongeren in toenemende mate problematisch te maken. Bovengenoemde gedragingen van leerlingen als het voortijdig de school verlaten, spijbelen en ordeverstoringen in de les, kunnen worden geïnterpreteerd als oplossingen voor deze spanningsvolle en onzekere situatie. Het zijn enkele mogelijkheden via welke leerlingen omgaan met hun onderwijs. Andere mogelijkheden zijn bijvoorbeeld het onderwijs gebruiken als parkeermogelijkheid met het oog op uitstel van werkloosheid en de mogelijkheid tot ontmoetingen met leeftijdsgenoten (Van Calcar & Bijlsma, 1985) of de mogelijkheid van een ongebreidelde jacht op diploma's en prestaties (Du Bois-Reymond, Karsten, Meijers, Rupp, Tillekens & Wendrich, 1985). Verschillende auteurs veronderstellen dat leerlingen, onder invloed van de genoemde maatschappelijke ontwikkelingen en de aanpassingen van het onderwijs daarop, geheel nieuwe betekenissystemen tot ontwikkeling brengen zoals bijvoorbeeld een leren met oog op de (toekomstige) vrije tijdsbesteding of een leren met oog op persoonlijke ontplooiing (Arbeitsgruppe, 1980; Van Gelder & Tillekens, 1986). In het onderzoek naar het scala van betekenisverleningen en gedragingen die leerlingen ontwikkelen in de omgang met hun onderwijs hanteren wij het begrip *leerlingperspectieven*. Dit onderzoek is in Nederland nog summier en weinig systematisch. In aansluiting op deze beperking stellen wij in het onderzoek onder meer de vraag: Welke perspectieven op onderwijs komen voor bij leerlingen uit het voortgezet onderwijs?

Het begrip perspectief gebruiken wij om de wijze waarop leerlingen omgaan met hun onderwijs te begrijpen en te beschrijven. Het is middels perspectieven dat mensen hun realiteit construeren en situaties definiëren. In navolging van Becker, Geer, Hughes & Strauss (1961) gebruiken wij het begrip perspectief 'to refer to a coordinated set of ideas and actions a person uses in dealing with some problematic situations, to refer to a person's ordinary way of thinking about and acting in such a situation. These thoughts and actions are coordinated in the sense that the actions flow reasonably from the actors point of view from the ideas contained in the perspective'. In hun definitie leggen Becker et al. (1961, noot 3) expliciet een relatie tussen opvattingen en handelen. Wij willen deze relatie niet op voorhand veronderstellen, maar de mogelijkheid behouden de relatie tussen betekenisverlening en handelen te bestuderen. Daartoe onderscheiden wij met betrekking tot het perspectief van leerlingen een betekeniscomponent en een handelingscomponent.

Met betrekking tot de betekeniscomponent achten wij een aantal domeinen van onderwijs relevant voor het perspectief van leerlingen: de relatie school-toekomst, het formele onderwijs-aanbod, het informele onderwijsgebeuren en de relatie van school tot het buitenschools leven.

Voor de illustratie van de analyse-stappen maken wij gebruik van de informatie die is verkregen over het domein 'het formele onderwijsaanbod'. Bij de beschrijving van de achtergronden van het onderzoek is nader uitgewerkt wat hieronder wordt begrepen.

Verschillende auteurs suggereren dat met de afnemende betekenis van het onderwijs voor een plaats op de arbeidsmarkt, leerlingen meer belang zullen gaan hechten aan actuele bruikbaarheid van de op school aan de orde gestelde leeractiviteiten. Daarnaast wordt gewezen op de toenemende zelfstandigheid van jongeren in buitenschoolse leefverbanden hetgeen in school, met de daar gehanteerde werkvormen, tot problemen zou resulteren. Onderzoeksgegevens aangaande betekenisverleningen van leerlingen aan hun onderwijs, ondersteunen deze veronderstellingen omtrent relevante aspecten van het formele onderwijsaanbod voor de perspectieven van leerlingen op hun onderwijs. Uit onderzoek valt af te leiden dat leerlingen met name weerstanden ervaren tegen de beperkte actuele relevantie van leerinhouden en tegen de beperkte initiatiefruimte die hun in lessen wordt geboden (Matthijssen, 1986).

Aan deze uitwerking van de achtergronden en het conceptueel raamwerk van het onderzoek kunnen enkele centrale vragen en sensitizing

concepts worden ontleend waarmee de analyse van het materiaal kan worden gestart:

Als sensitizing concepts worden gehanteerd:
1. Onderwijsperspectieven;
2. De relatie school-toekomst;
3. Het formele onderwijsaanbod (o.a. relevantie van de lesstof);
4. Het informele onderwijsgebeuren;
5. De relatie schools en buitenschools leren;
6. Handelingspatronen die voortvloeien uit de perspectieven.

Vragen die de analyse sturen:
1. Welke perspectieven hanteren leerlingen van het Voortgezet Onderwijs ten aanzien van de relatie school-toekomst?
2. Welke perspectieven hanteren leerlingen van het Voortgezet Onderwijs ten aanzien van het formeel onderwijsaanbod?
3. Welke perspectieven hanteren leerlingen van het Voortgezet Onderwijs ten aanzien van het informeel onderwijsgebeuren?
4. Welke perspectieven hanteren leerlingen van het Voortgezet Onderwijs ten aanzien van de relatie schools-buitenschools leren?
5. Welke perspectieven hanteren leerlingen van het Voortgezet Onderwijs ten aanzien van de samenhang tussen 1 t/m 4?
6. Hoe gaan leerlingen van het Voortgezet Onderwijs om met de schoolse leervoorwaarden?
7. Welke handelingspatronen vloeien uit de perspectieven voort?

Ook zijn er keuzen gemaakt omtrent waar en op welke wijze data worden verzameld:

Uit de probleemstelling en het conceptueel raamwerk valt af te leiden dat relevante schoolse voorwaarden voor de ontwikkeling van perspectieven op onderwijs door leerlingen zijn gelegen in ten eerste de marktwaarde van het diploma (hetgeen is gerelateerd aan het opleidingstype) en ten tweede de structuur van het curriculum (benoemd als het socialisatieregime). Overeenkomstig deze relevant geachte schoolse voorwaarden zijn scholen voor het onderzoek geselecteerd (zie Jongeling, 1986a). In eerste instantie zijn leerlingen van drie verschillende scholen in het onderzoek betrokken. Van elke school zijn 12 leerlingen geselecteerd: totaal 36 leerlingen.

Bij de dataverzameling is gebruik gemaakt van twee, elkaar aanvullende, informatiebronnen te weten observaties en interviews: het principe van data-triangulatie (Denzin, 1971). De informatie uit de observaties is in de gesprekken met de leerlingen ingebracht. Hun is gevraagd het gedrag onder de betreffende omstandigheden nader toe te lichten. Hiermee wordt enerzijds beoogd informatie te verkrijgen over de handelingscomponent van het onderwijsperspectief. Anderzijds heeft het tot doel de betekeniscomponent van het onderwijsperspectief op geldige wijze uit te werken door de alledaagse kennis van leerlingen over hun onderwijs te achterhalen (vgl. Jongeling, 1986b).

In het onderzoek wordt verondersteld dat bepaalde domeinen van het onderwijs een centrale plaats innemen in de betekenisverlening door leerlingen aan hun onderwijs. In de gesprekken met de leerlingen zijn diverse topics aan de orde gesteld als uitwerking van die domeinen. Onder andere met betrekking tot het formele onderwijsaanbod zijn de volgende topics gehanteerd:
* opvattingen ten aanzien van de lesstof waaronder opvattingen ten aanzien van actuele bruikbaarheid van het geleerde;
* opvattingen ten aanzien van de gehanteerde werkvormen;
* opvattingen ten aanzien van het optreden van leerkrachten;
* de relatie van deze aspecten tot de eigen belangen en behoeften;
* gezelligheid in de les;
* het eigen gedrag in diverse lessen.

Ook hieraan zijn ingangen ontleend om het verzamelde materiaal te analyseren.

2.2 *Stap 2: formuleren van voorlopige begrippen*

Nu de achtergronden, het conceptueel raamwerk, de onderzoeksvragen en de wijze van dataverzameling zijn onderzocht op relevante ingangen kan het materiaal worden geanalyseerd. Het gaat er bij de analyse in de exploratiefase om zo veel mogelijk aspecten te vinden die relevant zijn voor de onderzoeksvragen. Dit gebeurt middels het toekennen van 'begrippen' in de vorm van trefwoorden[2] aan tekstprotocollen. Met behulp van de volgende handelingen wordt dit bereikt.
a. Transcriptie van ruw materiaal;
b. Open coderen;
c. Reflecteren over de coderingen, conceptueel raamwerk en onderzoeksvragen in memo's.

De interviews worden eerst met behulp van een tekstverwerkingsprogramma uitgetypt. Vervolgens worden enkele interviews gekozen voor de eerste analysewerkzaamheden. In de exploratiefase is het niet nodig alle interviews in de analyse te betrekken; het doel is primair

een zo breed mogelijk scala aan betekenissen te selecteren, opdat de te ontwikkelen begrippen een breed geldigheidsbereik bezitten:

> De selectie van interviews heeft plaats gevonden nadat deze globaal zijn doorgelezen en wel zodanig dat van elke school vier leerlingen zijn gekozen die variëren in hun opvattingen ten aanzien van hun onderwijs. In totaal zijn dus twaalf interviews geselecteerd.

Om het materiaal geschikt te maken voor analyse wordt transcriptie toegepast. Transcriptie wil zeggen het indelen van ruw materiaal in hanteerbare en betekenisvolle eenheden: scènes. Zo is het mogelijk om elke nieuwe vraag en antwoord in een interview af te bakenen als een scène. In het onderhavige onderzoek wordt een scène gedefinieerd als een samenhangend geheel van uitspraken van een leerling ten aanzien van een interviewtopic.

In dit onderzoek is bij de analyse gebruik gemaakt van het computerprogramma KWALITAN[3]. KWALITAN biedt de mogelijkheid tegelijk met het afbakenen van scènes aan die scène reeds een eerste trefwoord toe te voegen. Van die mogelijkheid is in het onderzoek gebruik gemaakt:

> Aan de scènes die slaan op het formele onderwijsaanbod zijn trefwoorden toegevoegd als 'lesstof', 'lesvorm', met het doel het onderwerp van de scène te identificeren. Deze eerste trefwoorden corresponderen in belangrijke mate met de topics van de vraaggesprekken met leerlingen.

Het afgebakende tekstfragment, alsmede het daaraan toegevoegde trefwoord, is vervolgens met behulp van KWALITAN in de computer opgeslagen. Op dit moment is het materiaal 'analyse-klaar'.

De analyse in de exploratiefase noemt men open coderen: "The unrestricted coding of the data" (Strauss, 1987, p. 28). Doel van het open coderen is: "...to open up the inquiry. Every interpretation at this point is tentative" (ibid., p. 29). Het materiaal wordt regel voor regel doorgelezen, waarbij de onderzoeker tracht zoveel mogelijk ingangen te vinden bij de onderzoeksvragen. In de praktijk betekent dit het toekennen van allerlei begrippen in de vorm van trefwoorden aan de scènes. Dat kunnen de eerder geformuleerde sensitizing concepts zijn, maar ook kunnen het nieuwe begrippen zijn die geformuleerd worden naar aanleiding van de confrontatie met het materiaal. Dat deze begrippen nogal kunnen verschillen wat betreft hun analytisch niveau of theoretische lading is, gezien de doelstelling van het open coderen, niet relevant. In de specificatiefase wordt dat verder uitgewerkt.

> In het onderzoek zijn tijdens het open coderen een veelheid verschillende trefwoorden aan de scènes toegekend. Deze zijn vervolgens met behulp van de computer aan het scènebestand toegevoegd. Een voorbeeld van de manier waarop het een en ander is opgeslagen, is hieronder weergegeven:
>
> dokumentnummer : 3
> dokumentkode : 0143
> respondent: HAVO-V; Marjo; vrl; 4e klas; ned
>
> scènenummer : 20
> trefwoorden : *lesvorm
> groep: samenwerken
> groepwerk: leuk
> groepwerk: meebeslissen
>
> 1. I: Hoe vind jij dat groepswerk dat jullie hebben?
> 2. M: Ik vind het best wel goed dat ze dat doen, want je leert zo ook in een groep samen te werken.
> 3. En het is ook wel leuk om te doen.
> 4. I: Leuker dan klassikaal?
> 5. M: Ja, dat denk ik toch wel, want ik denk ook dat als je het klassikaal doet dan zeg ik veel minder, dan heb je veel minder inspraak en als je het in de groep doet ja, dan doe je het echt samen.
> 6. Ik denk wel dat dat scheelt.
> 7. Als je het klassikaal doet zijn er altijd een paar die wat zeggen en een paar die bijna nooit iets zeggen.
> 8. Maar als je het in de groep doet, beslist iedereen wel mee."

Op de analyse in de exploratiefase wordt gereflecteerd in memo's. 'Memoing' is een constant proces dat begint bij de eerste begrippen en doorgaat in de opeenvolgende fases van de analyse. In theoretische memo's worden begrippen ontwikkeld naar een hoger conceptueel niveau. In deze fase van het onderzoek is daarvan, gezien de doelstelling van het open coderen, nog geen sprake. Wel vindt een reflectie plaats over de trefwoorden die aan de scènes zijn toegevoegd. Daartoe vraagt de onderzoeker, met behulp van de computer,

een overzicht op van deze trefwoorden. Een voorbeeld van zo'n reflectie is hieronder gegeven:

> Het blijkt dat in de scènes die zijn geordend onder een hoofdtrefwoord als bijvoorbeeld lesvorm, uitspraken van leerlingen voorkomen aangaande verschillende aspecten van dat hoofdtrefwoord. Zo kunnen we ten aanzien van het hoofdtrefwoord LESVORM zes aspecten onderscheiden:
> 1. het samen met andere leerlingen werken aan een opdracht;
> 2. de aan leerlingen geboden initiatiefruimte in de les;
> 3. de structuur van het lesrooster;
> 4. de zeggenschap over de leerinhouden;
> 5. de wijze van verwerving van de leerstof;
> 6. de behoefte aan ontspanning in de les.

In methodische memo's reflecteert de onderzoeker op het conceptueel raamwerk en de onderzoeksvragen. Deze reflectie levert onder meer onderzoeksvragen voor nadere analyse van het materiaal. Hieronder is de vraag weergegeven voor nadere analyse van het tweede aspect van het hoofdtrefwoord LESVORM, de aan leerlingen geboden initiatiefruimte in de les:

> Welke opvattingen komen we bij leerlingen tegen ten aanzien van de hun geboden initiatiefruimte in de les?

De theoretische en methodische reflectie op het open coderen wordt afgesloten met recapitulatiememo's waarin de aangescherpte onderzoeksvragen en de meest centrale begrippen staan opgesomd, waarna de analyse verder gaat volgens de stappen in de specificatiefase. Voordat de analyse in de specificatiefase wordt besproken zullen we nader ingaan op enkele aspecten van het reflecteren tijdens het open coderen.

Het proces van open coderen is in principe een weg zonder einde. Altijd kan men nieuwe ingangen vinden om trefwoorden te formuleren of nieuwe vragen te stellen aan het materiaal. Als men steeds nieuw materiaal verzamelt betekent dit dat het proces van open coderen eindeloos voortgezet kan worden. (In feite zal de onderzoeker in volgende fases steeds ook met nieuw materiaal geconfronteerd worden en dan inderdaad de mogelijkheid hebben om nieuwe ingangen te formuleren of om de bestaande trefwoorden alsnog te veranderen.) Kortom, vanuit het empirisch materiaal zal er geen aanleiding zijn om de exploratiefase af te sluiten. Als men maar lang genoeg doorgaat in deze fase loopt het open coderen uit op een zee van trefwoorden waarin de onderzoeker dreigt te verzuipen. Vandaar dat in de memo's voortdurend wordt gereflecteerd op de vraag wat centrale aspecten zijn om de onderzoeksvraag te kunnen beantwoorden. Deze vraag tracht men mede te beantwoorden door de eigen bevindingen met de onderzoeksliteratuur te confronteren. Heeft de onderzoeker in deze zin greep gekregen op het trefwoordenbestand en geeft de analyse van nieuwe gevallen geen aanleiding om aan deze voorlopige invulling te twijfelen, dan gaat de onderzoeker over naar de volgende fase waarin de voorlopig geformuleerde begrippen nader worden uitgewerkt.

3 *Stappen in de specificatiefase*

Het eindmemo uit de exploratiefase levert een overzicht op van alle relevante trefwoorden die in die exploratiefase zijn gebruikt. Deze trefwoorden hebben een verschillende rol in het analytisch kader; sommigen staan centraler in de analyse dan anderen en nemen de vorm aan van theoretische dimensies of variabelen, anderen kunnen worden beschouwd als categorieën van variabelen, terwijl weer anderen te zien zijn als waarnemingsitems (bv. gespreksonderwerpen) of indicatoren. Al deze trefwoorden zijn tot nu toe beschouwd als voorlopige begrippen. In de specificatiefase wordt uitgezocht hoe al deze trefwoorden kunnen worden gerangschikt. Zo zal straks blijken dat sommige trefwoorden van hoog theoretisch niveau zijn (dimensies, variabelen), terwijl anderen van een lager theoretisch niveau kunnen worden beschouwd (categorieën, indicatoren). Door middel van de procedure van constante vergelijking, het voortdurend vergelijken van scènes waarbij wordt gelet op verschillen en overeenkomsten tussen de scènes, worden de trefwoorden geordend en uitgewerkt. Het vergelijken geschiedt met behulp van expliciete codeerprocedures. De theoretische ideeën die uit de vergelijkingen resulteren, worden geïntegreerd in de theorie-in-ontwikkeling. In de stappen is dit idee van constant vergelijken verder uitgewerkt.

3.1 *Stap 1: coderen van nieuw materiaal*

De eerste stap in de specificatiefase is het coderen van nieuw materiaal. Dit is een aan te bevelen tussenstap, omdat de begrippen uit de exploratiefase vaak worden ontwikkeld op een klein databestand. Het bestand wordt nu uitgebreid met nieuw materiaal (hier: meerdere interviews). Via toepassing van de begrippen op nieuw materiaal krijgt de onderzoeker meer duidelijkheid over de inhoud en reikwijdte van die begrippen. Tijdens het coderen van nieuw materiaal kunnen begrippen worden geherformuleerd en ook nieuwe begrippen worden geformuleerd. Deze stap houdt dus een controle in van de begrippen zoals tot nu toe gehanteerd. De onderzoekshandelingen tijdens deze stap zijn:

a. Keuze van nieuwe gevallen, nieuw materiaal;
b. Transcriptie van nieuw materiaal;
c. Coderen van materiaal;
d. Reflecteren op het coderen in memo's.

Eerst wordt bepaald welke eenheden vervolgens bij de analyse moeten worden betrokken en eventueel wordt dus nieuw materiaal verzameld (waarbij de exploratiefase ook aanpassing van het waarnemingsinstrument tot gevolg zou kunnen hebben). Als men een fase van verzameling invoert dan worden weer een aantal interviews of observaties gekozen om te worden geanalyseerd (vgl. 2.1). Ook is het mogelijk materiaal te kiezen uit een bestand dat reeds is verzameld. De volgende handeling is transcriptie van het materiaal. In 2.2 is beschreven wat transcriptie inhoudt en op welke wijze dit in het onderzoek is uitgevoerd. Nadat transcriptie heeft plaatsgevonden, gaat de onderzoeker coderen. Daartoe heeft deze nu een lijst met begrippen uit de exploratiefase ter beschikking. Scène voor scène wordt nagegaan of die begrippen 'passen' bij de tekst. Ook wordt bekeken of nieuwe elementen uit de scènes naar voren komen. Deze worden weergegeven in de vorm van nieuwe trefwoorden. In memo's wordt het een en ander verantwoord.

In het onderzoek heeft dit coderen van nieuw materiaal niet plaatsgevonden. De reden daarvoor is dat in de exploratiefase een intensieve analyse van 12 interviews heeft plaatsgevonden. Verwacht werd dat de relevante aspecten van het veld van onderzoek daarmee zijn gevonden.

3.2 *Stap 2: vergelijking van scènes van hetzelfde begrip*

In de tweede stap start de onderzoeker met het ordenen, afbakenen en uitwerken van de begrippen en daarmee met de constante vergelijking van scènes. Tijdens deze stap bepaalt de onderzoeker welke rol de voorlopige begrippen hebben in het analytisch kader. Daartoe worden scènes van een bepaald centraal trefwoord met elkaar vergeleken. Daarbij wordt de vraag gesteld welke andere trefwoorden te zien zijn als aspecten van dat centrale trefwoord. Strauss (1987, p. 32) noemt dit proces 'axial coding'. We proberen nu dus de trefwoorden die bij elkaar horen te ordenen in groepjes trefwoorden (variabelen, indicatoren etcetera). De volgende onderzoekshandelingen worden tijdens deze stap verricht:

a. Scènes selecteren;
b. Vergelijken van geselecteerde scènes per begrip;
c. Reflecteren op de vergelijkingen in memo's (clusters trefwoorden formuleren).

Voorwerk voor het uitvoeren van vergelijkingen tijdens deze stap is het selecteren van scènes. De onderzoeksvragen uit de exploratiefase zijn daarvoor een belangrijke ingang.

In het onderhavig onderzoek zijn bijvoorbeeld zes selecties gemaakt van scènes die betrekking hebben op de zes onderscheiden aspecten van het hoofdtrefwoord LESVORM (zie 2.2).

De scènes zijn vervolgens per selectie met elkaar vergeleken. Dit komt er in de praktijk op neer, dat men tijdens intensief doorlezen van de scènes let op verschillen en overeenkomsten daartussen. Daarbij is de aandacht gericht op die informatie die relevant is voor het onderscheiden aspect; andere informatie in de scène wordt (voorlopig) buiten beschouwing gelaten.

Het resultaat van de vergelijkingen wordt weergegeven in memo's. Ten eerste hebben de vergelijkingen groepjes indicatoren opgeleverd die betrekking hebben op aspecten van lesvorm. Hieronder zijn de groepjes trefwoorden met betrekking tot enkele aspecten van lesvorm weergegeven:

1. Samenwerken met andere leerlingen
   groep: leren samenwerken
   groep: elkaar kunnen helpen

groep: gezelligheid
groep: samenstelling
groep: afhankelijkheid
groep: leergemak
groep: onvoldoende discipline
groep: irrelevant
groep: leuk
liefst alleen werken

6. De behoefte aan ontspanning in de les
werken + ontspanning
leren>ontspanning
ontspanning bevordert concentratie
ontspanning>leren

Ten tweede hebben de vergelijkingen enkele hypothesen opgeleverd omtrent de dimensies rond een begrip waarnaar zo'n groepje indicatoren verwijst. Een voorbeeld hiervan is:

> Hypothese 3: een groep scènes onder lesvorm heeft betrekking op verschillende waarderingen van leerlingen van de in de les bestaande initiatiefruimte. Die waarderingen verwijzen naar meerdere dimensies. Het lijkt daarbij te gaan om: de behoefte aan ontspanning in de les, de ervaren leeropbrengst van zelfstandig werken en de bereidheid tot zelfdiscipline.

Met deze voorlopige bevindingen rond een hoofdtrefwoord gaat de onderzoeker in de volgende stap verder.

3.3 *Stap 3: vergelijking van scènes van clusters begrippen onderling*

Tijdens deze stap worden door middel van vergelijkingen de cluster van trefwoorden rond een begrip uitgewerkt naar een begrip van een hoger conceptueel niveau. Vaak nemen zulke begrippen de vorm aan van dimensies. De dimensies worden zoveel mogelijk geformuleerd in termen van twee uiteinden van de dimensie (Peters et al., 1989; Strauss, 1987). Voordeel daarvan is dat de dimensie sterk wordt afgebakend (vgl. Glaser & Strauss, 1965, 'de besefcontext': 'open besefcontext' versus 'gesloten besefcontext'). Belangrijke ingangen voor het op deze wijze uitwerken van begrippen zijn de hypothesen die tijdens de vorige stap zijn geformuleerd. Het uitwerken van dimensies geschiedt op de volgende wijze:

a. Selecties van scènes op neventrefwoorden van een centraal begrip;
b. Vergelijken van geselecteerde scènes;
c. Reflecteren op de vergelijkingen in memo's (dimensies formuleren);
d. Tegenhangers formuleren;
e. Definiëren van begrippen (dimensies en indicatoren).

Voorwerk voor het vergelijken van de scènes is wederom de selectie van scènes, nu aan de hand van de geformuleerde hypothesen. De geselecteerde scènes worden door middel van intensief doorlezen met elkaar vergeleken op verschillen en overeenkomsten. In memo's worden de resultaten van de vergelijkingen beschreven en worden dimensies gedefinieerd. De inhoudelijke resultaten van de vergelijkingen met betrekking tot dimensies omtrent de waardering van initiatiefruimte (zie hypothese), worden hieronder weergegeven:

> *De waardering van initiatiefruimte*
> Vergelijkingen van relevante scènes aan de hand van de hypothese hebben meerdere dimensies opgeleverd.
>
> De eerste dimensie heeft betrekking op de *behoefte aan ontspanning* bij de uitvoer van leeractiviteiten in de les: de 'ontspanningsbehoefte'. Leerlingen variëren in die behoefte en waarderen op grond daarvan lessituaties verschillend. Twee polen van de dimensie zijn: *'ontspannen leren'* en *'efficiënt leren'*.
>
> Bij *ontspannen leren* krijgen de leeractiviteiten voor leerlingen betekenis vanuit de behoefte aan ontspanning bij de uitvoer daarvan, zonder dat het ten koste gaat van het leren: een evenwichtige afwisseling tussen leren en ontspannen wordt nagestreefd. Ontspanning wordt gewaardeerd omdat het bijdraagt tot een betere concentratie op de leeractiviteiten. Klassikale lessen worden in deze visie als bedreigend voor de ontspanning ervaren (saai, langdradig, onvrij). Leerlingen hebben in dit soort lessen dan ook moeite om 'bij de les te blijven'. Lesvormen met meer initiatiefruimte voor leerlingen worden positief gewaardeerd, vanwege de vrijheid die het biedt tot ontspanning in de les en tot verdeling van de werklast over school en thuis.
>
> Bij *efficiënt leren* krijgen lessituaties voor leerlingen betekenis vanuit de behoefte zoveel mogelijk leeractiviteiten binnen school te voltooien. Hierbij wordt de behoefte aan ontspanning ondergeschikt gemaakt aan de wens tot leren. School wordt door deze leerlingen gedefinieerd als leerinstituut. De motivatie voor dit accent op leren is onder andere dat deze leerlingen de buitenschoolse tijd zoveel mogelijk wensen te beschouwen als vrij te besteden tijd. De waardering van initiatiefruimte in de lessen is genuanceerd: zij wordt afhankelijk gemaakt van de verhouding tussen de te leveren inzet en

ervaren leeropbrengst bij zelfwerkzaamheid. Groepswerk wordt daarin, evenals klassikale lessen, negatief gewaardeerd omdat de afhankelijkheid van anderen (leerlingen en leerkracht) voor hulp en uitleg als te groot wordt ervaren. Individueel werken wordt daarentegen positief gewaardeerd: het biedt de mogelijkheid in eigen tempo te werken.

Een tweede dimensie met betrekking tot de waardering van initiatiefruimte slaat op samenwerking tussen de leerlingen en betreft de ervaren leeropbrengst van *zelfstandig (individueel of in groepsverband) werken*. Twee polen van de dimensie zijn *'zelfstandigheid leren'* en *'leren is consumeren'*.

Bij *zelfstandigheid leren* wordt de initiatiefruimte gewaardeerd vanuit opbrengsten als de bijdrage tot het leren zelfstandig en onder eigen verantwoordelijkheid te werken en de bijdrage tot het leren samenwerken met anderen. Klassikale lessen worden zo negatief gewaardeerd, omdat in die lessen slechts een passief consumeren van leerlingen wordt verwacht.

Bij *leren is consumeren* wordt door leerlingen bij de evaluatie van lesvormen waarbij zelfwerkzaamheid wordt verwacht, op geen enkele wijze gerefereerd aan leeropbrengsten als boven beschreven. Individueel werken en groepswerk worden negatief gewaardeerd terwijl klassikale lessen en de daarbij horende passieve houding worden gewaardeerd.

Een derde en laatste dimensie met betrekking tot de waardering van initiatiefruimte slaat op het orde-aspect en houdt verband met de *bereidheid tot zelfdiscipline*. *'Leren onder leiding van de leerkracht'* en *'leren onder eigen beheer'* zijn als polen te herkennen.

Bij *leren onder leiding van de leerkracht* worden lesvormen waarin leerlingen zelf de verantwoordelijkheid dragen voor de planning en organisatie van leeractiviteiten door hen afgewezen. Men verwijst daarbij voornamelijk naar de wanorde die deze lessen met zich mee zouden brengen. Lespatronen als beurten en klassegesprekken worden positief gewaardeerd, omdat de leerling daarbij gedwongen wordt 'bij de les te blijven'.

Bij *leren onder eigen beheer* wordt juist die eigen verantwoordelijkheid ten aanzien van planning en organisatie van leeractiviteiten geaccepteerd.

Tot zover de beschrijving van de in het onderzoek geformuleerde dimensies. In een volgende stap in de specificatiefase worden de dimensies uitgewerkt naar variabelen. In de derde fase van de g.t-benadering, de reductiefase, wordt de samenhang tussen die variabelen onderzocht. Deze stappen worden in de volgende paragraaf besproken als mogelijke stappen voor verder onderzoek.

4 *Stappen voor verder onderzoek*

Tot nu toe zijn in het onderzoek via verschillende stappen theoretische dimensies geformuleerd. Om de analyse verder richting te geven worden aan het eind van de specificatiefase variabelen geconstrueerd vanuit die dimensies. De ontwikkeling van de dimensies heeft namelijk plaatsgevonden op basis van vergelijkingen van *scènes*. De variabelen echter hebben betrekking op de *eenheden*. Zoals de dimensies zijn weergegeven, in termen van twee uiteinden op een continuüm, is nog niets bekend over de uiteindelijke positie van een respondent op dat continuüm. Het is mogelijk dat verschillende respondenten variëren tussen de posities op de dimensie. Bij de constructie van de variabelen worden die variaties bepaald en herleid tot posities op de variabele. Om dit te bewerkstelligen verandert de onderzoeker van analyse-niveau: in de analyse staan niet langer scènes centraal, maar de eenheden (hier: de leerlingen). Bij deze stap komt een nieuw element in de kwalitatieve analyse naar voren: het gebruik van overzichtskaarten (Miles & Huberman, 1984; Wester, 1987; Peters et al., 1989)[4]. Met behulp van vergelijkingen op overzichtskaarten wordt het mogelijk variabelen te construeren en de kwalitatieve score per eenheid op die variabelen te bepalen. Vanuit de scores van de eenheden op deze variabelen kunnen verschillende typen eenheden (leerlingen) zichtbaar worden.

De overstap naar de volgende fase in de g.t.-benadering, de reductiefase, wordt gevormd door hypothesen omtrent samenhangen tussen centrale variabelen. Het onderzoek naar samenhangen tussen de variabelen is eveneens gebaseerd op het principe van constante vergelijking, waarbij de overzichtskaart weer een belangrijk hulpmiddel is. Het doel van de analyses in de reductiefase is het formuleren van de kern van de te ontwikkelen theorie. Wat de centrale optiek zal worden om het geheel te ordenen is nu nog niet duidelijk. Het kan gaan om een theoretisch construct of een basisproces (bv. identiteitsontwikkeling) dat de relaties tussen de centrale variabelen goed thematiseert. Een andere mogelijkheid is

dat een kernbegrip kan worden geformuleerd bijvoorbeeld op basis van een bundeling van verschillende, op één of andere wijze samenhangende, variabelen. Typerend voor een kernbegrip is dat het centraal staat in het onderzoek, het meest met alle andere begrippen samenhangt. Er zijn verschillende strategieën beschikbaar om zo tot een kernbegrip te komen (zie hiervoor: Peters et al., 1989). Zowel voor een kernbegrip als voor een theoretisch construct of basisproces geldt dat de beslissing om daarmee de centrale optiek in te vullen via literatuurstudie wordt onderbouwd. In het geïllustreerde onderzoek zal de centrale optiek betrekking hebben op perspectieven van leerlingen ten aanzien van het onderwijsaanbod.

In de laatste fase van de g.t.-benadering, de integratiefase, worden vanuit de centrale optiek relaties gelegd met de centrale begrippen in het onderzoek zoals het socialisatieregime, de maatschappelijke positie van leerlingen en het opleidingstype. Met behulp van kwalitatieve analyses (via de overzichtskaart), dan wel kwantitatieve analyses (variantie-analyse) kunnen de verschillende relaties nauwkeurig worden onderzocht. Meestal is daarvoor aanvullende materiaalverzameling vereist (zie Denzin 1987).

5 *Discussie*

In dit artikel hebben we aandacht geschonken aan het analyseproces in de eerste twee fasen van een kwalitatief onderzoek. De voorgestelde systematisering en concretisering van het analyseproces levert de onderzoeker een aantal voordelen op.

Ten eerste wordt het mogelijk doelgericht te analyseren. Door aan het proces van gefundeerde theorieontwikkeling fasen te onderscheiden en per fase zowel de doelstellingen voor analyse als de rol van te ontwikkelen begrippen daarbij te expliciteren, krijgt de onderzoeker meer zicht op het analyseproces als geheel. Het expliciete karakter van de stappen stelt de onderzoeker tevens in staat continu na te gaan waar deze zich bevindt in het proces van analyse.

Ten tweede zijn de, in de handboeken vaak algemeen getypeerde, analyseprocedures geconcretiseerd naar onderzoekshandelingen. Aan de onderzoeker wordt duidelijk gemaakt hoe de analyse-procedures in de praktijk van onderzoek gestalte krijgen als onderdeel van het proces van kwalitatieve analyse.

Ten derde wordt het proces van analyse beter over te dragen en te verantwoorden. Keuzen en beslissingen kunnen nu duidelijk worden geëxpliciteerd. Dit is ten eerste van belang voor de communicatie binnen een onderzoeksteam, of met de begeleidingscommissie van een onderzoek. De diverse betrokkenen beschikken nu over een gemeenschappelijk referentiekader voor wat betreft de methodiek van analyse van het onderzoeksmateriaal. Daarnaast kan het bijdragen aan de evaluatie van de kwaliteit van de resultaten van kwalitatief onderzoek. Kortom, de hier gepresenteerde werkwijze kan gezien worden als een bijdrage aan de verbetering van de reproduceerbaarheid van de kwalitatieve analyse.

In deze bijdrage is met name ingegaan op het analyseproces in de eerste twee fasen van het onderzoek. Dit heeft tot gevolg dat diverse aspecten van de kwaliteit van kwalitatief onderzoek hier niet aan de orde hebben kunnen komen (voor een uitgebreide bespreking zie o.a. Smaling, 1987; Wester, 1987). Op een enkel punt willen we hier ten slotte kort ingaan, omdat het direct met het analyseproces verbonden is.

Ten aanzien van de start van een kwalitatief onderzoek wordt veelal gedacht dat de onderzoeker verplicht is om met zo weinig mogelijk (theoretische dan wel inhoudelijke) voorkennis het onderzoeksveld te betreden. Dit is een misverstand. Een goed uitgewerkte *probleemstelling* staat geenszins op gespannen voet met een open benadering[5]. De open benadering houdt met name in dat deze probleemstelling en het analytisch kader dat daarbij is gehanteerd, voortdurend aangepast kunnen worden op grond van naar voren komende inzichten. De cyclus reflectie-waarneming-analyse-reflectie zorgt voor voortdurende impulsen in dit proces. Een van de bronnen waaruit dit proces gevoed wordt is de inhoudelijke en theoretische *literatuur*. Kortom, literatuurstudie ten behoeve van en voorafgaand aan de formulering van de probleemstelling wordt niet als afgerond beschouwd en loopt tijdens het gehele onderzoeksproces door. In ieder geval geldt dat de relevantie van hetgeen aan de literatuur wordt ontleend empirisch moet worden onderbouwd. Het bovenstaande neemt niet weg dat

men in 'Discovery' wordt aangeraden om als men *niet* op een bepaald veld thuis is, eerst het veld te verkennen en dan pas uitgebreid literatuur te raadplegen. Immers, de daar naar voren gebrachte inzichten zijn beter toegankelijk als men ze kan verbinden met de praktijken waarnaar zij verwijzen. Maar meestal doet men onderzoek op een veld waar men (de onderzoeker of de onderzoeksgroep waarin men participeert) *wel* thuis is. Er zijn dan eerder te veel dan te weinig ideeën vooraf. Glaser en Strauss hebben 'Discovery' geschreven vanuit het perspectief van de creatieve onderzoeker die in discussie met het empirisch materiaal delen van deze voorkennis tot onderbouwde inzichten kan ontwikkelen. De kwalitatieve onderzoeker is dus voortdurend bezig het onderzoeksprobleem te *theoretiseren*. Het is duidelijk dat met name ervaren onderzoekers eerder over de daarvoor noodzakelijk kwaliteiten zullen beschikken dan minder ervaren onderzoekers.

Noten

1. Onder begrippen worden verstaan alle termen die een rol spelen in het analytisch kader-in-ontwikkeling: dus hypothesen, relaties, theoretische begrippen, aspecten, eigenschappen, variabelen, indicatoren, topics, categorieën.
2. Met de term 'trefwoorden' wordt verwezen naar de praktische analysewerkzaamheden van het toevoegen van begrippen aan tekstprotocollen. De term 'begrippen' wordt gebruikt in relatie tot de theoretische en methodische reflectie over de trefwoorden.
3. Tijdens het onderzoek heeft het programma KWALITAN een puur administratieve functie. Het materiaal moet tijdens de kwalitatieve analyse namelijk meerdere malen worden geselecteerd en geordend. Dit brengt een administratieve rompslomp met zich mee waarin relatief veel onderzoekstijd gaat zitten. Met behulp van KWALITAN tracht de onderzoeker aan de behoefte van administratieve ondersteuning te voldoen. Voor een beschrijving van de mogelijkheden van KWALITAN verwijzen we naar Peters en Wester, 1988 en Peters et. al., 1989.
4. Het valt buiten het bereik van dit artikel nader in te gaan op het gebruik van de overzichtskaart in kwalitatieve analyse. Voor bijzonderheden daarover verwijzen we naar de aangehaalde auteurs.
5. Sommigen (bv. Glaser & Strauss, 1967) hanteren een open probleemstelling, dat wil zeggen zo weinig mogelijk literatuurstudie vooraf. Anderen (m.n. Miles & Huberman, 1984; Maso, 1987) benadrukken juist het belang van een sterk op literatuur gerichte fase vooraf aan het onderzoek om zodoende tot een uitgewerkte probleemstelling te komen. Weer anderen (Wester, 1987; Strauss, 1987) nemen een tussenpositie in; literatuurstudie vooraf is noodzakelijk om zodoende richtinggevende en attenderende ('sensitizing') begrippen en eventueel de samenhang daartussen te kunnen formuleren. Deze begrippen geven, zeker in de eerste analysefase, richting aan de analyse.

Literatuur

Arbeitsgruppe Schulforschung, *Leistung und Versagen. Alltagstheorien von Schülern und Lehrern*. München: Juventa, 1980.

Becker, H.S., B. Geer, E.C. Hughes & A.L. Strauss, *Boys in white*. Chicago: University of Chicago Press, 1961.

Bernards, Y., Spijbelen, schakel tussen school en vrije tijd. In: *Jeugdonderzoek in Nederland*. Programmaboek congresjeugdonderzoek in Nederland. Rijksuniversiteit Utrecht: 1986, p. 34-51.

Bois-Reymond, M. du, S. Karsten, F. Meijers, J. Rupp, G. Tillekens & E. Wendrich, De krisis van de school en de krisis in de onderwijswetenschappen. *Comenius*, 1985, nr. 18.

Bois-Reymond, M. du & T. Veld, Jongeren en hun toekomst: over de krisis in het onderwijs in de jaren tachtig. In: G. Tillekens (red.), *Het opgelegde leren*. Nijmegen: ITS, OOMO-reeks, deel 3, 1986.

Calcar, C. van & K. Bijlsma-Frankema, Schooluitval, feiten en verklaringen. In: *Onderwijs met perspectief op arbeid*. 's-Gravenhage: SVO/COD, 1985.

Denzin, N.K., *The research act: A theoretical introduction to sociological methods*. Chicago: Aldine, 1978.

Gelder, B. van & G. Tillekens, De herzuiling van het onderwijs. In: G. Tillekens, *Het opgelegde leren*. Nijmegen: ITS, OOMO-reeks, deel 3, 1986.

Glaser, B.G. & A.L. Strauss, *The discovery of grounded theory; strategies for qualitative research*. New York: Aldine, 1967.

Glaser, B.G., *Theoretical sensitivity; Advances in the methodology of grounded theory*. Mill Valley: Sociology Press, 1978.

Jongeling, E.E.M., HAVO-leerlingen over hun onderwijs. In: M..A.J.M. Matthijssen, W. Meeus & F. van Wel (red.), *Beelden van jeugd*. Groningen: Wolters-Noordhoff, 1986a.

Jongeling, E. E. M., Methodische overwegingen bij onderzoek naar schoolbeleving en gedrag van HAVO-leerlingen. In: P. Grob, R. Hüsken & P. van der Kley (red.), *Het opgelegde leren*. Nijmegen: ITS, OOMO-reeks, deel 4, 1986b.

Matthijssen, M. A. J. M., *De ware aard van balen*. Groningen: Wolters-Noordhoff, 1986.

Maso, I., *Kwalitatief onderzoek*. Meppel: Boom, 1987.

Miles, M. B. & A. M. Huberman, *Qualitative data analysis. A sourcebook of new methods*. Beverly Hills: Sage, 1984.

Peters, V., *KWALITAN, een ondersteunend programma bij de analyse van kwalitatieve gegevens*. Nijmegen: Vakgroep Methoden Maatschappijwetenschappen, Katholieke Universiteit Nijmegen, Nijmegen: 1987.

Peters, V., F. Wester, De computer bij kwalitatief onderzoek. Een onderzoeksnotitie. *Sociologische Gids*, 1988, *35*, 332-345.

Peters, V., F. Wester & R. Richardson, *Kwalitatieve analyse in de praktijk en handleiding bij Kwalitan versie 2*. Nijmegen: ITS, SWI-reeks, 1989.

Smaling, A., *Methodologische objectiviteit en kwalitatief onderzoek*. Lisse: Swets & Zeitlinger, 1987.

Strauss, A. L., *Qualitative Analysis For Social Scientists*. Cambridge: Cambridge University Press, 1987.

Wester, F. P. J., *De gefundeerde theoriebenadering* (dissertatie). Nijmegen: 1984.

Wester, F., *Strategieën voor kwalitatief onderzoek*. Muiderberg: Coutinho, 1987.

Curricula vitae

R. Richardson studeerde methoden en technieken van sociaal onderzoek aan het Sociologisch Instituut van de K.U. te Nijmegen. Was daar twee jaar werkzaam op het gebied van het systematiseren van het analyseproces in kwalitatief onderzoek. Is momenteel werkzaam bij het Instituut voor Sociale Pedagogiek en Andragogiek aan de K.U. te Nijmegen, alwaar hij kwalitatief onderzoek verricht naar het intern functioneren van ouderenorganisaties.

E. Jongeling studeerde onderwijssociologie aan de K.U. te Nijmegen. Was werkzaam als toegevoegd onderzoeker bij de vakgroep Socialisatie en Ontwikkeling aan de R.U. te Utrecht. Heeft daar kwalitatief onderzoek verricht naar schoolmotivatie van leerlingen in het voortgezet onderwijs.

F. Wester is wetenschappelijk medewerker bij de vakgroep Methoden van Maatschappijwetenschappen aan de K.U. te Nijmegen. Zijn belangstelling gaat uit naar de methodologie van het kwalitatief onderzoek. Enige publikaties: "De Gefundeerde Theorie-benadering", dissertatie, Nijmegen, 1984; "Strategieën voor kwalitatief onderzoek", Muiderberg, 1987; en samen met W. Arts en H. Hilhorst "Betekenis en Interactie. Het symbolisch interactionisme als onderzoeksperspectief", Deventer, 1985.

Adres: Vakgroep Algemene Pedagogiek, Katholieke Universiteit Nijmegen, Postbus 9103, 6500 HD Nijmegen

Manuscript aanvaard 19-2-'90

Summary

Richardson, R., E. Jongeling & F. Wester. 'Systematic analyses in the grounded theory; an illustration.' *Pedagogische Studiën*, 1990, *67*, 355-367.

Despite methodological guide-lines for qualitative research methods, still there is a lack of clear prescriptions for this type of research. In doing qualitative research, therefore, researchers meet with a lot of problems and pitfalls. The authors of this article tried to solve this problem in relation to one of the most applied strategies of qualitative research; the grounded theory-approach. They describe some steps of analysis as directives for systematic analysis in the grounded theory-approach. The computerprogram KWALITAN is an important help in carrying out these steps of analysis. The steps are illustrated with data from the researchproject: "Pupil perspectives on education in Secondary Schools in Holland".

Kinderen en autonomie in de vrije tijd

B. A. J. VAN DER WOUW en
J. W. TE KLOEZE
Landbouwuniversiteit Wageningen

Samenvatting

Vrijetijdsbesteding in onze samenleving vindt meestentijds plaats in huiselijke sfeer of met huisgenoten. Kinderen zijn voor de inrichting van hun vrije tijd voor een groot deel afhankelijk van toestemming en goedkeuring van hun ouders. In dit onderzoek (onder 331 moeders) wordt een meetinstrument gepresenteerd voor de autonomie die ouders verlenen bij de inrichting van de vrije tijd van hun kinderen. In de socialisatie-literatuur wordt ouderlijke controle gezien als een belangrijk aspect van een opvoedingsstijl: een door de tijd relatief stabiele manier van omgang tussen ouders en kinderen. Het door Kohn gepresenteerde theoretische model ter verklaring van verschillen in opvoedingsstijlen is aan de hand van beschikbare gegevens getoetst. In tegenstelling tot Kohns suggestie zijn geen directe verbanden gevonden tussen milieu-variabelen en autonomieverlening. Wel zijn sterke verbanden gevonden tussen opvoedingswaarden en autonomieverlening. Ook gezinsstructurele variabelen blijken van invloed op autonomieverlening. Binnen autonomieverlening met betrekking tot vrije tijd kan een onderscheid gemaakt worden naar controle over tijd (wanneer een kind iets mag) en controle over activiteiten (of een kind iets mag). Moeders blijken het meest restrictief ten aanzien van het eerste: de vraag wanneer iets mag.

1 *Inleiding*

De relatie tussen ouders en hun kinderen is een relatie die het best getypeerd kan worden als een opvoedingsrelatie. Daarmee is een structurele asymmetrie in die relatie ingebouwd. In het verleden is dan ook een grote hoeveelheid onderzoek verricht naar relaties tussen ouderkenmerken en opvoedingsgedrag aan de ene kant en de sociale en persoonlijkheidsontwikkeling van het kind aan de andere. Een recent overzicht van met name Amerikaans onderzoek vanuit dit perspectief geven Peterson en Rollins (Peterson & Rollins, 1987). We signaleren de laatste jaren een toenemende kritiek op de zogenaamde 'social mold' theorieën, die ervan uitgaan dat opvoeding een unidirectioneel proces is. Opvoeding, zo wordt gesteld, is juist een bidirectioneel proces waarin ook het kind een actieve rol speelt (Peterson & Rollins, 1987; Walters & Stinnett, 1971) en tevens een proces dat niet los gezien kan worden van de verdere gezins- en sociale context. Het is echter duidelijk dat in de relatie ouder-kind grote verschillen in macht en competentie aanwezig zijn. De zeggingskracht van de ouder is zeker voor een groot deel bepalend te achten voor hetgeen plaatsvindt in de context van het gezin. Dit betekent dat de mate waarin de ouders hun zeggenschap naar zich toe trekken of juist overdragen naar een kind een belangrijke invloed zal hebben op de aard van de interacties binnen de ouder-kind relatie. In dit artikel gaan we in op de vraag in hoeverre de variatie in autonomieverlening zijn verklaring vindt in de variatie in opvoedingswaarden, milieukenmerken en gezinsstructurele kenmerken.

2 *Opvoedingsstijlen, -gedrag en -waarden*

Veel onderzoek op het terrein van de ouder-kindrelatie houdt zich bezig met het identificeren van kenmerken waarin ouders in hun gedragingen en houdingen ten opzichte van hun kinderen substantieel van elkaar zijn te onderscheiden (Maccoby & Martin, 1983). Men is op zoek naar opvoedingsstijlen. We kunnen het begrip opvoedingsstijl in navolging van Vervoort definiëren als: "De wijze waarop de ouders zich met betrekking tot vragen van vrijheid, zelfstandigheid, onafhankelijkheid e.d. tegenover de kinderen opstellen" (Vervoort, 1968, p. 131). Een opvoedingsstijl is ons inziens daarmee ook een door de tijd relatief stabiele manier van omgang tussen ouders en hun kinderen. Analytisch is een op-

Pedagogische Studiën 1990 (67) 368-381

voedingsstijl uiteen te leggen in een gedragscomponent en een waardencomponent. De gedragscomponent verwijst naar hoe de ouders zich feitelijk (zullen) gedragen c.q. autonomie verlenen, de waardencomponent naar hoe de ouders zelfstandigheid, vrijheid etc. in hun kinderen waarderen. Het meeste onderzoek met betrekking tot opvoedingsstijlen richt zich vrijwel uitsluitend op de waardencomponent. In de meeste gevallen worden typeringen van opvoedingsstijlen gehaald uit de dimensies die naar voren komen uit factoranalyses van vragen aan ouders over gewaardeerde eigenschappen of gedragingen van kinderen (Maccoby & Martin, 1983). Over het algemeen worden volgens deze methode twee belangrijke dimensies binnen een opvoedingsstijl onderkend:

a. *ouderlijke ondersteuning* met als polen warmte/affectie en vijandigheid/afstandelijkheid
b. *ouderlijke controle* met als polen permissief en restrictief (Maccoby & Martin, 1983; Peterson & Rollins, 1987; Rollins & Thomas, 1979).

Ouderlijke controle is in onderzoek op velerlei wijzen geoperationaliseerd. Soms als specifiek beïnvloedingsgedrag (bijv. het gebruik van fysieke kracht, dreigen met ontzegging van liefde), soms via de effecten van ouderlijk gedrag op het kind, soms ook als een algemeen gezinsklimaat (of rolstructuur). Rollins en Thomas (1979) geven in hun literatuuroverzicht alleen al 18 verschillende benamingen die voor ouderlijke controle worden gebruikt. Bovendien lijkt ouderlijke controle meer te verwijzen naar de *manier* waarop ouders invloed proberen uit te oefenen (dwang, dreigen met onthouding van liefde, redeneren) en niet naar uit te oefenen controle in termen van *de mate van* feitelijk te verlenen autonomie.

Dat ten aanzien van ouderlijke controle geen eenduidigheid in operationalisering bestaat is ook niet verwonderlijk. Zelfstandigheid van kinderen en hoe ouders daar tegen aankijken kan immers uiteenlopende betekenissen hebben. Het leren op straat op eigen benen te kunnen staan tegenover leeftijdsgenoten is een ander soort zelfstandigheid dan je keurig kunnen gedragen in een restaurant of adequaat telefonische boodschappen kunnen overbrengen (Newson & Newson, 1976). Vaak wordt zelfstandigheid in verband gebracht met het in meerdere of mindere mate egalitair zijn van de ouder-kind relatie. Het betreft dan bijvoorbeeld de vraag in hoeverre ouders hun kinderen laten meepraten en meebeslissen over gezinsaangelegenheden. Maar het kan ook verwijzen naar de mate waarin ouders hun kinderen de vrije hand geven bij het nemen van beslissingen over hun eigen doen en laten. Bijvoorbeeld hoe vrij ouders hun kinderen laten in de bepaling en invulling van hun eigen vrije tijd.

Kinderen vertoeven het merendeel van hun tijd onder toezicht van volwassenen (ouders, leerkrachten etc.). Maar er zijn wel delen van de dag waarop zij zich aan dit toezicht ten dele of geheel kunnen onttrekken. Op weg van school naar huis bijvoorbeeld. Of je terugtrekken op je eigen kamer. Wanneer je bij een vriendje of vriendinnetje huiswerk gaat maken of gaat logeren. Dit betekent dat ouderlijke controle ook kan worden geoperationaliseerd als het min of meer zelfstandig beslissingen laten nemen over deze 'vrijetijdsdomeinen'. Het leek ons dan ook zinvol een instrument te ontwikkelen dat juist op het punt van autonomieverlening ingrijpt, een instrument waarin variabelen opgenomen zijn die verwijzen naar een aantal belangrijke beslissingsmomenten voor de inrichting van de vrije tijd. We willen daarbij nagaan welke mate van autonomie ouders hun kinderen verlenen op verschillende (soorten) domeinen van de vrije tijd.

3 *Opvoeding en vrije tijd*

In Nederland hebben Kamphorst en Spruijt (1983) onderzoek verricht naar het belang van socialisatieprocessen voor vrijetijdsgedrag. Op basis van door hen bestudeerde literatuur maken ze een onderscheid tussen socialisatie tot vrijetijd(sgedrag) en socialisatie via vrijetijd(sgedrag). Zij onderstrepen het belang van vrije tijd voor het gezinsleven. Facetten van persoonlijkheidsvorming als karaktervorming, sociabiliteit, coöperatie en ontwikkeling van de eigen identiteit komen in onderzoek naar voren als 'pedagogisch produkt' van omgaan met vrije tijd. Zij concluderen op basis van eigen onderzoek dat dit omgaan met vrije tijd een sterke samenhang vertoont met socialisatievariabelen uit het ouderlijk gezin. Wat er in de vrije tijd in de gezinssfeer wordt

gedaan en hoe daar mee wordt omgesprongen, is volgens hen meer bepaald door de ervaringen in het ouderlijk gezin, dan de huidige sociale determinanten als opleiding en sociale klasse. In het verlengde hiervan stellen zij dat in het socialisatie-onderzoek teveel de nadruk wordt gelegd op milieuspecifieke socialisatie en te weinig op de seksespecifieke socialisatie.

Kamphorst en Spruijt hebben onderzocht of opvoedingsstijlen een samenhang vertonen met de participatie in vrijetijdsgedrag als geheel. Zij hebben gekeken in hoeverre een opvoedingsstijl het vrijetijdsgedrag op latere leeftijd beïnvloedt. Een sterk autoritaire opvoedingsstijl blijkt een wat sociabeler vrijetijdsgedrag op latere leeftijd te bewerkstelligen. Een matig autoritaire opvoeding leidt tot een opstelling waarbij mensen wat meer moeite voor dingen doen en wat minder consumptief zijn. Een wat zij noemen sterk kindgerichte opvoedingsstijl maakt mensen creatiever en ondernemender. Een matig kindgerichte opvoeding resulteert in een minder 'zorgzaam huiselijk' vrijetijdsgedrag. De gevonden relaties tussen opvoedingsstijl en vrijetijdsgedrag zijn bij kindgerichte stijlen sterker dan bij autoritaire.

Het belang van de relatie tussen opvoeding en vrije tijd tonen ook Van der Heiden en Smeding. Zij laten zien dat vrije tijd een belangrijk domein is waarop gezinsopvoeding aangrijpt. Ouders stimuleren uit pedagogische overwegingen, maar ook met het oog op gezinsintegratie of ook wel uit plichtsgevoel het samen iets ondernemen met het kind. Er zijn ook gezinnen waarin men weinig samen doet omdat men dat niet nodig vindt, of om juist het belang van zelfstandigheid van de kinderen (zelfredzaamheid, zelf vermaken) te benadrukken. Verschillen wat betreft dit samen doen lijken samen te hangen met gezinsfase (Van der Heiden & Smeding, z.j.). In gezinnen met heel jonge kinderen èn in gezinnen met oudere kinderen is de mate waarin men dingen samen doet geringer dan in gezinnen met opgroeiende kinderen. Tevens bemoeilijken grote leeftijdsverschillen tussen de kinderen en – in mindere mate – het hebben van veel kinderen gemeenschappelijke activiteiten.

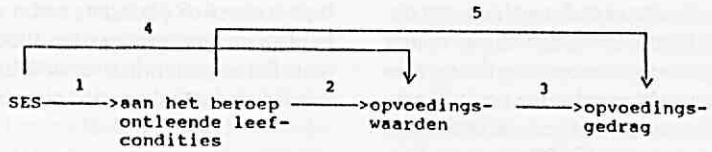
4 *Opvoedingsstijlen, hun milieu- en seksegebondenheid*

Vooral met betrekking tot de klassebepaaldheid van opvoedingsstijlen is een veelheid aan literatuur verschenen (voor een overzicht, zie bijv. Gecas, 1979). Een van de meest populaire onderwerpen daarbij is het onderzoek naar de 'modes of parental control'. Alhoewel de bevindingen niet altijd even consistent zijn, wordt door Gecas een aantal generalisaties opgevoerd dat door de onderzoeksliteratuur ondersteund wordt. Belangrijk voor dit betoog zijn de bevindingen dat in de hogere sociale milieus de ouder-kindrelatie meer gelijkwaardig is en minder autocratisch dan in de lagere sociale milieus en dat zelfstandigheid eveneens meer wordt benadrukt in de hogere sociale milieus.

Godsdienstigheid (de geïnvolveerdheid van een gezin in een religieuze cultuur), is veel minder uitgebreid onderzocht op zijn effecten voor opvoeding. Wel wordt het belang van godsdienst en geloofsbeleving (dus ongeacht de denominatie) in een gezin voor attitudes ten aanzien van zaken als huwelijk, man-vrouw-verhouding en socialisatie allerwegen onderkend (vgl. bijv. Kooy, 1975). In een onderzoek van Knijn en Verheijen onder Nijmeegse moeders van jonge kinderen bleek kerksheid een belangrijke indicator voor traditionaliteit ten aanzien van moederschapsopvattingen (Knijn & Verheijen, 1986).

Het meest uitgewerkte theoretische kader voor de relatie sociaal milieu en opvoedingsstijl is afkomstig van Melvin Kohn (zie Gecas, 1979). Zijn veronderstelling is dat de leefsfeer van de ouders sterk van invloed is op het door hen gehanteerde waardensysteem. Iemands ervaringen in met name de beroepssfeer zijn bepalend te achten voor de ontwikkelde waardenoriëntaties op het terrein van (in dit geval) opvoeding. Deze waardenoriëntaties leiden het handelen van ouders ten opzichte van hun kinderen.

Hoofdarbeiders zullen, doordat hun levensomstandigheden een hogere graad van zelfbepaling toestaan c.q. vereisen, in de opvoedingsrelatie meer nadruk leggen op zelfbepalingswaarden (zelfstandigheid, individualiteit, initiatief, creativiteit, zelfverwerkelijking), terwijl de levensomstandigheden van handarbeiders meer op conformiteit en autoriteit georiënteerd zijn, waardoor zij con-



Figuur 1 *Model van Kohn (ontleend aan Gecas, 1979)*

formiteitswaarden (orde, netheid, gehoorzaamheid) in de opvoeding zullen benadrukken. Ook in Nederland is aandacht besteed aan de relatie tussen sociaal milieu en opvoedingswaarden in de zin van zelfbepaling versus conformiteit (Meijnen, 1979, 1984; Jansen, 1983; Bosker, Meijnen & Van der Velden, 1986). Steeds weer blijkt een belangrijke samenhang gevonden te worden tussen sociaal milieu en opvoedingswaarden. Overigens dient aangetekend te worden dat daarbij niet alleen de eigenschappen van het beroep, maar veel meer opleidingsniveau de belangrijkste voorspeller blijkt te zijn van iemands waardenoriëntaties (Bosker e.a., 1986; Peterson & Rollins, 1987). In het door Kohn gehanteerde verklaringsmodel worden verbanden gesuggereerd, die in Figuur 1 zijn weergegeven.

Vooral de linker helft van het model is empirisch veelvuldig onderbouwd. Veel minder is dat het geval met de relaties 3 en 5. Gecas (1979) geeft aan dat Kohn in het model sociaal milieu zowel direct als indirect relateert aan ouderlijk opvoedingsgedrag alleen *omdat* er zo'n sterke relatie is tussen sociaal milieu en opvoedingswaarden. Empirische onderbouwing ontbreekt echter. De veronderstelling in het model is dat wanneer ouders nadruk op zelfstandigheidswaarden leggen dit zal uitmonden in opvoedingsgedrag waarin het kind een grotere mate van zelfstandigheid, en daarmee zeggingskracht krijgt toegewezen. Alhoewel er nooit sprake is van geobserveerd gedrag, maar altijd van geverbaliseerd (intentioneel) gedrag, vinden we in de literatuur aanwijzingen dat hier de relatie met sociaal milieu wel eens veel minder prominent zou kunnen zijn dan de relatie tussen sociaal milieu en zelfbepalingswaarden. In recent Iers onderzoek (Hynes, 1985) kon geen milieuspecifiek verband worden aangetoond tussen zelfbepaling als opvoedingswaarde en bestraffend opvoedingsgedrag. Newson en Newson (1976) tonen in hun onderzoek onder zevenjarige kinderen eerder een verschil tussen seksen aan dan een klasseverschil, vooral wanneer het gaat om het feitelijk verlenen van handelingsvrijheid. Ondersteuning voor het sekse-specifieke karakter van disciplinering in de opvoedingsrelatie vinden we ook in het onderzoek van Jansen (1983) naar conformiteits- en zelfbepalingswaarden. Het lijkt ook alsof de wijze waarop ouders verschillen tussen jongens en meisjes percipiëren hun opvoedingsgedrag beïnvloedt. "Specifically working mothers and parents with nontraditional sex-role attitudes tend to encourage greater independence, autonomy and achievement orientations in female offspring" (Peterson & Rollins, 1987, p. 486).

En in een interactionistisch georiënteerde Nederlandse studie naar de opvoedingsreacties van ouders op overtredingsgedrag van kinderen, is gebleken dat objectieve kenmerken van de opvoedingssituatie (waaronder geslacht en leeftijd van de overtreder en de aard en gevolgen van de overtreding) veel sterker met het opvoedingsgedrag samenhangen dan milieukenmerken als beroeps- en opleidingsniveau van de ouders (Gerris & Janssens, 1987). Ouders blijken hun opvoedingsgedrag sterk af te stemmen op de kenmerken van de situatie (wat is er aan de hand?), alsook op de kenmerken van de opvoedeling (over wie gaat het?). Tot de situationele kenmerken zouden naast de leeftijd en geslacht van de opvoedeling ook algemene gezinsstucturele kenmerken (als gezinsgrootte en sekse-verdeling van de kinderen) gerekend kunnen worden. De vraag is of de situationele kenmerken diezelfde dominante rol spelen wanneer opvoedingsgedrag op andere wijze dan als bestraffingsgedrag geoperationaliseerd wordt.

5 *Probleemstelling*

Zoals gezegd, gaat het om opvoedingsgedrag van ouders ten opzichte van hun kinderen en de opvoedingswaarden die zij daarbij hante-

ren, beide deel uitmakend van een opvoedingsstijl. Wij beschouwen daarbij de relatie tussen gedrag en waarden zodanig dat opvoedingswaarden geacht worden het opvoedingsgedrag te beïnvloeden. Opvoedingsgedrag (i.c. ouderlijke controle) heeft in ons geval betrekking op de mate van feitelijk te verlenen autonomie door ouders aan kinderen voor de inrichting van de vrije tijd.

De centrale probleemstelling omvat allereerst de vraag naar de mate van autonomieverlening door ouders aan hun kinderen op domeinen van vrije tijd, en de vraag of er binnen deze domeinen dimensies zijn aan te wijzen. Vervolgens vragen we ons af in hoeverre de mate van autonomieverlening afhankelijk is van opvoedingswaarden als zelfbepaling en sekse-onderscheid, en of ook de door Kohn veronderstelde relaties met sociaal milieu gevonden kunnen worden, wanneer godsdienstigheid en gezinssituationele variabelen in het verklaringsmodel worden opgenomen.

6 *Data*

De gegevens voor dit artikel zijn afkomstig uit een onderzoek naar gezinssituaties op het Nederlandse verstedelijkte platteland. In 1986 zijn daarvoor alle gehuwde vrouwen uit twee plattelandsplaatsen met tenminste één thuiswonend kind in de leeftijd van 0 tot 18 jaar benaderd. In totaal kwamen 547 gezinnen in aanmerking om te worden opgenomen in het onderzoek. Uiteindelijk zijn 331 verwerkbare interviews gehouden. Vergelijking van kenmerken van de responsgroep en de nonresponsgroep (leeftijdsverdelingen man en vrouw, autochtoniteit, huwelijksduur, huishoudensgrootte, gezinsfase) leverde geen noemenswaardige verschillen op. De gegevens zijn derhalve representatief te achten voor de twee plattelandsgemeenten, die op hun beurt weer model kunnen staan voor het Nederlandse verstedelijkte platteland (Douma, 1975).

De gemiddelde leeftijd van de vrouwen is 37 jaar. Dertig procent heeft een kind ouder dan vijftien jaar, de helft heeft al een kind in de tienerleeftijd, terwijl 11% nog géén schoolgaande kinderen heeft. Het aantal gezinnen met twee kinderen komt in meer dan de helft van de gevallen voor. 82% is niet geboren in een der beide plaatsen, maar komt van elders. Tevens is 80% van de echtgenoten geboren buiten de twee plaatsen, zodat we te maken hebben met een grote groep allochtone gezinnen. Dit is gezien het verstedelijkte karakter van het onderzoeksgebied niet verwonderlijk.

7 *Instrumenten*

Bij de constructie van de vragenlijst is gezocht naar een aantal vragen die ouders iets laten zeggen over hoe zij feitelijk omgaan met hun kinderen als het gaat om beslissingen over de indeling en invulling van hun vrije tijd. Het gaat om de invloed die ouders zeggen uit te (zullen) oefenen op het gedrag van hun kinderen. Een zevenentwintigtal items maakt deel uit van de ontwikkelde schaal voor AUTONOMIEVERLENING. Deze schaal bestaat uit negen vragen over alledaagse zaken die spelen bij de opvoeding van kinderen, zoals het tijdstip waarop het kind naar bed gaat, hoe het kind zijn eigen kamer inricht, welke tv-programma's het kind mag kijken etc. Het betreft zaken die naar ons oordeel in voldoende mate die vrijetijdsdomeinen dekken waar het autonomieprobleem tussen ouders en kinderen kan spelen. De onderscheiden items betreffen de keuzevrijheid inzake domeinen die direct of indirect op vrije tijd of vrijetijdsactiviteiten betrekking hebben. Aangezien te verwachten is dat ouders sterk naar de leeftijd van een kind zullen kijken bij autonomieverlening is steeds gevraagd voor de leeftijden 8, 12 en 16 jaar aan te geven, in hoeverre de ouders de beslissingsmacht over die zaken naar zichzelf toe trekken, of anders gezegd in welke mate zij hun kind hierin vrij laten. Aangezien moeders zijn ondervraagd met kinderen tussen 0 en 18 jaar, is het mogelijk dat zij ouder zijn van kinderen die wel respectievelijk niet 8, 12 of 16 jaar zijn. Derhalve is de vraagformulering als volgt geweest: "Wilt u het antwoord aangeven waarvan u denkt dat het het dichtst komt bij zoals u het zou doen als uw kind(eren) zo oud is/zijn of zou(den) zijn?" We nemen daarbij voor lief dat ouders van wie de kinderen één of meer van die leeftijden zijn gepasseerd, retrospectief moeten antwoorden, terwijl ouders van kinderen die nog niet één of meer van die leeftijden hebben gehad, prospectief moeten antwoorden. In beide gevallen kan de situatie rooskleuriger worden voorgesteld (al dan niet op basis van sociaalwenselijk gedrag) dan die was of zal zijn.

De antwoordmogelijkheden liepen van 1 (Wij laten onze kinderen helemaal vrij in hun keuze) tot 4 (Dat beslissen wij als ouders helemaal zelf). De vier antwoordmogelijkheden indiceren niet onmiddellijk bepaalde handelingen (of het nalaten daarvan) maar potentiële en/of intentionele gedrags*componenten* (bijvoorbeeld in gesprekken, al dan niet gepland, of terloops, waarin wordt gezegd wat mag of wat niet mag, wat ouders verbieden of toelaten, met alle variaties daartussen). De vier antwoordmogelijkheden functioneren in dat opzicht als modellen (op het terrein waar het hier over gaat) voor opvoedingsgedrag, waarvan ouders aangeven of men zich daarin kan herkennen.

Voor de kenschetsing van het milieu waarin de gezinnen te plaatsen zijn, gebruiken we twee indices. Een index voor sociaal economische status (SES) van het gezin, waarvoor vier variabelen zijn gebruikt: het beroepsniveau en het opleidingsniveau van zowel man als vrouw. De variabele beroepsniveau is gemeten volgens de methode van het ITS (Westerlaak, Kropman & Collaris, 1975). De indeling van de variabele opleidingsniveau is overgenomen van het SCP. Via hercodering is gekomen tot een vierpuntsschaal lopend van 1. lager onderwijs met evt. LBO tot 4. HBO of Universiteit. De gehanteerde indicator voor sociaal economische status is gebaseerd op de door Bernstein en Brandis ontworpen en door Meijnen (1979) beschreven methode, uitgaande van het onderscheid hand- versus hoofdarbeid in beroepsniveau en de eerder genoemde vierdeling in opleidingsniveau. De tweede door ons gehanteerde index voor gezinsmilieu is godsdienstigheid (geloof), gemeten als een participatie-index (o.a. kerkgang, bidden, bijbellezen) bestaande uit 10 items. De alpha van deze schaal bedraagt 0.86 (zie Van der Wouw, nog te publiceren).

Op het terrein van opvoedingswaarden zijn twee schalen gebruikt. Voor de schaal ZELFBEPALING is gebruik gemaakt van een Nederlandse versie van de schaal van Kohn zoals die is uitgewerkt door Meijnen (1979) in zijn onderzoek naar de relatie tussen sociale klasse en intelligentie. Uit een lijst van 10 karaktereigenschappen is de moeders gevraagd welke drie zij voor een kind van 10 jaar het belangrijkste vonden en vervolgens deze te rangschikken. Evenzo voor de drie meest onbelangrijke eigenschappen. De resterende eigenschappen kregen de middenwaarde (4). De aan Meijnen ontleende scoringsprocedure levert een schaal op met een reikwijdte van 28 tot 52. Naarmate de score hoger is worden meer zelfbepalingswaarden benadrukt.

De tweede door ons gebruikte schaal ter indicering van opvoedingswaarden is SEKSEONDERSCHEID: een uit vijf items bestaande schaal waarin gevraagd wordt of men het eens is met een sekse-onderscheid tussen bepaalde vrijetijds-activiteiten als crossfietsen, met poppen spelen, aan brommers sleutelen, lezen, ruwe sporten (alpha = 0.83). Hoge scores indiceren een sterke afwijzing van dit sekse-onderscheid.

Als indicatoren voor gezinsstructuur zijn gebruikt GEZINSFASE (de leeftijd van het oudste kind), GEZINSGROOTTE en SEKSEVERDELING DER KINDEREN.

8 *Analyses en resultaten*

De frequentieverdelingen en gemiddelde scores op de 27 items voor autonomieverlening laten grote verschillen zien tussen de aan de orde gestelde zaken (Tabel 1).

De keuze van een sport wordt door bijna alle ouders vrijwel geheel aan hun kinderen overgelaten. Ook het inrichten van de slaapkamer, welke vriendjes of vriendinnetjes een kind te logeren mag vragen, wat een kind tijdens de vakantie mag doen en naar welke televisieprogramma's het mag kijken, zijn over het algemeen zaken waarin kinderen behoorlijk vrij worden gelaten. Minder vrijheid krijgt het kind als het erom gaat of het kind bij een vriendje of vriendinnetje mag gaan logeren. Sterk in eigen hand houden de ouders zaken als hoe laat een kind mag thuiskomen, hoe laat het naar bed gaat en of het kind wel of niet mee op vakantie gaat.

Alle gemiddelden blijken bij stijgende leeftijdscategorie te dalen, hetgeen impliceert dat inderdaad ouders hun kinderen bij stijgende leeftijd meer vrijheid geven. Het grootste verschil over de leeftijden zien we optreden bij het tijdstip van naar bed gaan en welk tv-programma kijken, gevolgd door het meegaan op vakantie. Het minst leeftijdsgevoelig zijn het te logeren vragen van een vriendje, de keuze van een sport en het gaan logeren bij een vriendje. De twee 'logeer'-items zijn wel de items waarover de moeders de meest uiteen-

Tabel 1 *Autonomieverlening, procentuele verdelingen van scores, gemiddelde scores en standaarddeviaties per variabele*

categorieën:		1. Wij laten onze kinderen helemaal vrij in hun keuze	2. Wij laten onze kinderen hierin behoorlijk vrij	3. Wij geven onze kinderen daarbij enige vrijheid	4.Dat beslissen wij als ouders helemaal zelf		gem.	st.dev.
variabelen:								
De keuze van een sport	8 jr	42	30	26	2		1.9	.87
	12 jr	54	35	11	0		1.6	.68
	16 jr	77	18	5	0		1.3	.54
						SPORT	1.6	.61
Hoe laat thuiskomen	8 jr		1	5	94		3.9	.28
	12 jr		2	31	67		3.7	.51
	16 jr	5	30	45	20		2.8	.82
						THUISKOM	3.5	.42
Hoe laat naar bed	8 jr		0	5	95		4.0	.24
	12 jr	1	4	36	59		3.5	.59
	16 jr	20	35	34	11		2.4	.92
						BEDGAAN	3.3	.47
Mee op vakantie	8 jr		1	5	95		3.9	.26
	12 jr	1	3	12	84		3.8	.51
	16 jr	19	28	37	16		2.5	.96
						VAKANMEE	3.4	.46
Logeren bij vriend(in)	8 jr	19	17	29	35		2.8	1.12
	12 jr	21	24	33	21		2.5	1.05
	16 jr	33	35	26	6		2.0	.92
						GALOGER	2.5	.92
Welk TV-progr. kijken	8 jr	10	15	34	42		3.1	.97
	12 jr	16	30	44	10		2.5	.88
	16 jr	59	27	12	2		1.6	.76
						TVKIJK	2.4	.71
Hoe slaapkamer inrichten	8 jr	18	25	35	22		2.6	1.02
	12 jr	31	40	26	3		2.0	.83
	16 jr	67	26	6	1		1.4	.60
						KAMER	2.0	.69
Wie komt logeren	8 jr	21	22	33	25		2.6	1.08
	12 jr	24	31	33	13		2.3	.98
	16 jr	31	40	23	7		2.1	.90
						VRALOGER	2.3	.90
Wat doen tijdens vakantie	8 jr	12	24	42	22		2.7	.94
	12 jr	15	41	36	8		2.4	.84
	16 jr	43	44	12	1		1.7	.72
						VAKANDOE	2.3	.71

lopende meningen hebben (de grootste standaarddeviaties). De minste variantie in antwoorden vinden we bij het item hoe laat het kind mag thuiskomen.

Kinderen van 16 jaar worden in hun keuze hierbij door nog geen 5% van de ouders vrij gelaten; 20% geeft aan dat zij dat als ouders geheel zelf beslissen. De vragenbatterij blijkt met een alpha van Cronbach: 0.86 te voldoen aan de eisen voor schaalbaarheid (vgl. Swanborn, 1982). We kunnen daarom stellen dat ouders c.q. moeders een soort overall opvoedingsgedrag ten toon spreiden ten aanzien van (het groeien naar) zelfstandigheid van hun kinderen, dat kan variëren van zeer permissief tot zeer restrictief. Optelling van de scores op alle 27 items en deling door 27 levert de schaal: AUTONOMIEVERLENING. De schaal blijkt nagenoeg normaal verdeeld te zijn, hetgeen erop wijst dat restrictief en permissief opvoedingsgedrag de uiterste polen zijn van een continuüm. Hoge scores op deze schaal geven aan dat de ouders sterke invloed op beslissingen uitoefenen.

Hoewel in eerste instantie een-dimensionaliteit van het instrument aanwezig lijkt, is het desniettemin mogelijk dat er binnen de itemset subschalen aanwezig zijn. Met andere woorden: zijn er clusters van terreinen aan te wijzen waarover de ouders hun kinderen in meer of mindere mate zelf laten beslissen? Om deze vraag te beantwoorden zijn per terrein de drie scores (voor de leeftijden 8, 12 en 16 jaar) opgeteld en daarna aan een factoranalyse onderworpen (Tabel 2).

Wanneer we ook nog even terugkijken naar de gemiddelde scores op de onderscheiden items, kunnen we een onderscheid maken in terreinen waarover meestentijds de kinderen een aardig woordje mogen meespreken, naast terreinen die in eerste aanleg beslissingsgebieden van de ouders zijn. De factor VTDOEN is te interpreteren als de vrijheid die ouders geven waar het zaken betreft met vaak een incidenteler karakter, en handelend over de invulling van vrijetijds-activiteiten, waarbij de ouderlijke controlemogelijkheid direct of indirect (gaan logeren) aanwezig blijft. Het gaat er om òf een kind iets mag doen, terwijl de factor VTTIJD meer verwijst naar de zeggingskracht van ouder en kind over de vraag wannéér een kind iets mag doen. Het gaat dan om de tijd die kinderen krijgen toegemeten om iets te ondernemen waarop de ouders (of andere ouders) weinig greep hebben. Hiermee is het eerste deel van de probleemstelling behandeld.

Tabel 2 *Factorladingen terreinen van ouderlijke invloed*

	Factor 1 VTDOEN	Factor 2 VTTIJD
Doen in vakantie	.68	
Kamer inrichten	.68	
TV kijken	.64	
Gaan logeren	.63	
Te logeren vragen	.62	
Keuze sport	.37	
Thuiskomen		.81
Mee op vakantie		.81
Naar bed gaan		.75
% verkl. variantie	32.2	15.3

Teneinde het tweede onderdeel van de probleemstelling te kunnen beantwoorden, zijn allereerst voor de drie schalen afzonderlijk regressieanalyses uitgevoerd. Aan de hand van multiple regressie met stapsgewijze selectie van voorspellers is gezocht naar het relatieve belang van waarden, milieu en gezinsstructuur voor de voorspelling van opvoedingsgedrag. In Tabel 3 zijn de gestandaardiseerde bèta-waarden aangegeven voor de onderscheiden regressie-vergelijkingen. Significante bèta-waarden ($p.<0.5$) zijn onderstreept.

Het meest opmerkelijke hierbij is dat de milieukenmerken geen significante voorspellers blijken te zijn voor de mate van AUTONOMIEVERLENING. Wel blijken dat in sterke mate beide opvoedingsoriëntaties en in mindere mate de gezinsstructuurkenmerken gezinsgrootte en sekseverdeling der kinderen te zijn. Alleen bij de subschaal VTTIJD blijkt sociaal economische status een significante voorspeller. Dit wijst erop dat wat betreft beslissingen over de vraag wannéér een kind iets mag doen in de vrije tijd, hogere sociale milieus enigszins permissiever zijn. Het verkregen meetinstrument AUTONOMIEVERLENING blijkt onafhankelijk te zijn van het sociale milieu. Sterke verbanden vinden we wel met de opvoedingswaarden ZELFBEPALING en SEKSE-ONDERSCHEID. Deze beide opvoedingsoriëntaties zijn op hun beurt wel weer in belangrijke mate gerelateerd aan

Tabel 3 *Multiple regressie-analyses van schalen van autonomieverlening; weergegeven zijn de gestandaardiseerde bèta-coëfficiënten*

	AUTONOMIE-VERLENING Bèta	VTDOEN Bèta	VTTIJD Bèta
Milieu			
SES	0.08	0.05	0.12
Godsdienstigheid	0.05	0.01	0.00
Gezinsstructuur			
Gezinsgrootte	0.12	0.08	−0.05
Sekseverdeling	0.11	0.07	0.18
Gezinsfase	−0.06	−0.06	0.09
Opvoedingsoriëntaties			
Zelfbepaling	−0.21	−0.19	−0.25
Sekse-onderscheid	−0.23	−0.23	−0.15

het sociale milieu zoals uit Tabel 4 blijkt.

Tabel 4 *Correlaties tussen milieukenmerken en opvoedingsoriëntaties*

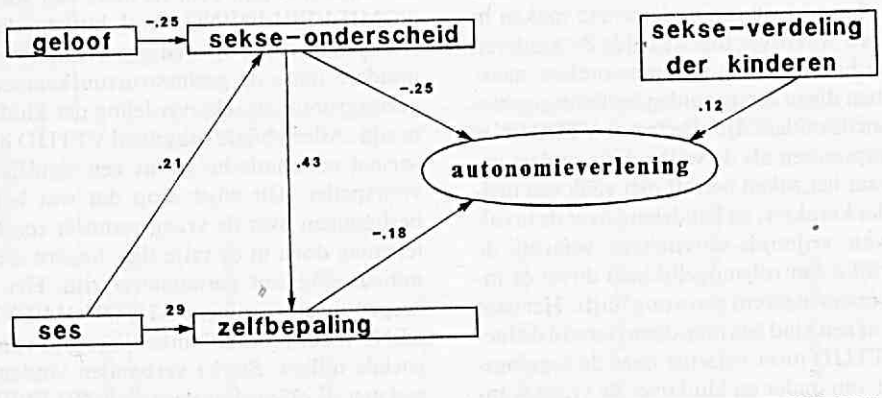
	ZELF-BEPALING	SEKSE-ONDERSCHEID
SES	0.37	0.19
Godsdienstigheid	−0.15	−0.23

Gepoogd is daarom met behulp van een padanalyse te komen tot een model, waarin milieu, gezinsstructuur, opvoedingswaarden en opvoedingsgedrag zijn opgenomen. De analyse is aangepakt met LISREL VI (Jöreskog & Sörbom, 1985). Het volgende theoretische model was daarbij uitgangspunt:

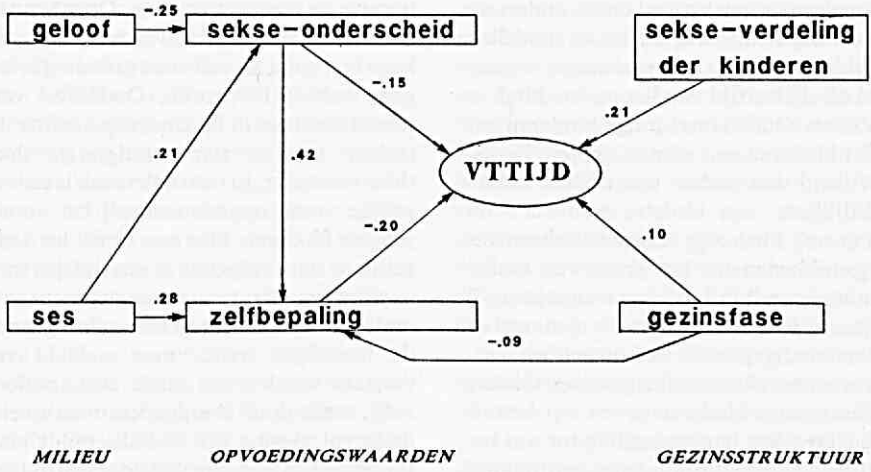
milieu → opvoedingsoriëntaties → opvoedingsgedrag ← gezinsstructuur

Uiteindelijk voldeed het volgende model (adjusted goodness of fit index = .954, chi-kwadraat = 10.52, p = .06) het best (Figuur 2).

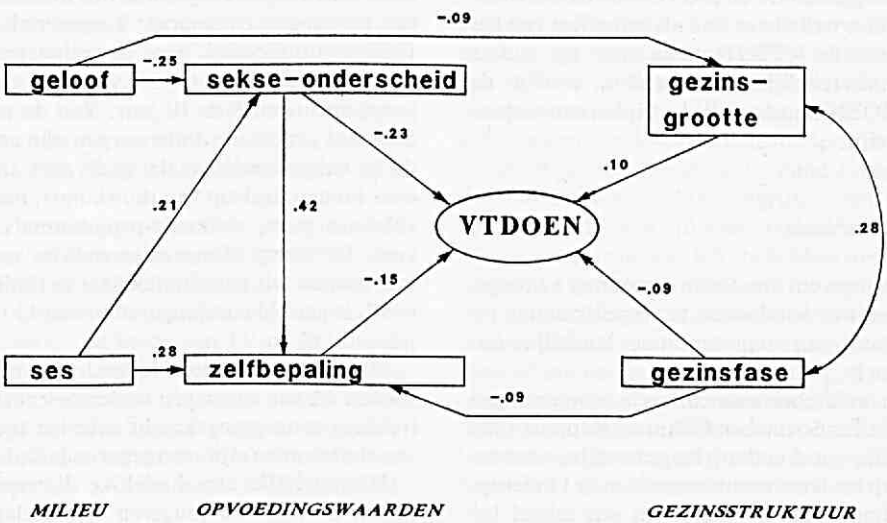
De mate waarin ouders beslissingen over de vrije tijd van hun kinderen naar zich toe trekken is in dit model afhankelijk van een tweetal ook onderling sterk samenhangende opvoedingswaarden, namelijk ideeën over hoe zelfstandig kinderen behoren te zijn en of zij



Figuur 2 *Padmodel van AUTONOMIEVERLENING*



Figuur 3 *Padmodel van*



Figuur 4 *Padmodel van VTDOEN*

in verband met de opvoeding een sterk onderscheid tussen jongens en meisjes afwijzen. Milieuvariabelen hebben alleen een indirecte invloed, religiositeit vooral via de opvoedingswaarde sekse-onderscheid, sociale status zowel via sekse-onderscheid als zelfbepaling. Een directe invloed heeft ook de gezinsstructuur. Ouders van alleen jongens laten hun kinderen het meest vrij in hun beslissingen, ouders van alleen meisjes laten hun kinderen het minst vrij. De invloed van sekse doet zich dus gevoelen zowel via de gehanteerde sekse-specifieke opvoedingswaarden, als wel via het situationele effect van de geslachtsverhouding onder de kinderen in het gezin.

Eenzelfde modelanalyse is gepleegd voor de twee subschalen VTTIJD en VTDOEN. De in de Figuren 3 en 4 weergegeven modellen leveren met name voor wat betreft de milieuvariabelen en de opvoedingswaarden globaal hetzelfde beeld. De invloed die uitgaat van de geslachtsverhouding onder de kinderen in een gezin blijkt zich met name te richten op wanneer een kind iets mag (Figuur 3 en Figuur 4).

Met betrekking tot de vraag of een kind iets mag, heeft het aantal kinderen in een gezin in-

vloed op de mate van invloed die de ouders op die beslissingen uitoefenen. In de modellen voor beide subschalen is de gezinsfase – geïndiceerd als de leeftijd van het oudste kind – opgenomen. Ouders met jonge kinderen geven hun kinderen een relatief grotere beslissingsvrijheid dan ouders met oudere kinderen. Blijkbaar zijn ouders, wanneer hun kinderen nog klein zijn, optimistischer over de mogelijkheden tot het geven van beslissingsruimte aan hun kinderen wanneer ze 12 of 16 jaar zijn, dan de ouders die reeds een of meerdere kinderen in die leeftijd hebben, tenminste waar het gaat om tijds*indeling*. Ouders met alleen jonge kinderen geven wat betreft vrijetijds*invulling* in tegenstelling tot wat betreft tijds*indeling* hun kinderen een relatief geringere beslissingsvrijheid dan ouders met oudere kinderen. Deze omkering in de relatie tussen gezinsfase en AUTONOMIEVERLENING is wellicht te zien als een effect van het feit dat de VTTIJD-items meer op oudere leeftijd feitelijk gaan spelen, terwijl de VTDOEN-items op alle leeftijden van toepassing zijn.

9 *Discussie*

Deze gegevens kunnen in een ruimer kader geplaatst worden door ze te vergelijken met resultaten van representatief landelijk onderzoek.

In onderzoek naar culturele veranderingen maakt het Sociaal en Cultureel Rapport 1986 melding van trends op lange termijn, waaronder op het terrein van 'vrijheid in de kinderopvoeding'. Aan de hand van een schaal bestaande uit drie items wordt vastgesteld dat er sinds 1965 sprake is van een toenemend liberalisme in de kinderopvoeding. Opvallend is dat deze trend sinds 1983 over haar hoogtepunt heen is, en dat nu zelfs een geleidelijke teruggang valt te bespeuren. Onderdeel van de schaal 'vrijheid in de kinderopvoeding' is het tijdstip waarop een twintigjarige dochter thuis moet zijn. In ons onderzoek is een soortgelijke vraag opgenomen, zij het voor wat jongere kinderen. Hiermee wordt het wel mogelijk de data enigszins in een tijdsperspectief te plaatsen. Wanneer we de gegevens uit ons onderzoek plaatsen tegen de achtergrond van de landelijke trend, mag wellicht verondersteld worden dat in de autonomieverlening, zoals door ons geoperationaliseerd, in de jaren zeventig een liberaliserende tendens te ontdekken viel, gevolgd door een duidelijke stabilisering in de jaren tachtig.

Een tweede landelijk onderzoek dat zich leent voor vergelijkingen met ons materiaal is het Nationaal Onderzoek Jongerenbladen, 1987, een onderzoek naar de activiteiten en bezits- en gebruikspatronen van kinderen en jongeren tussen 6 en 19 jaar. Van de negen door ons genoemde onderwerpen zijn er drie die op enigerlei wijze in dat onderzoek aan de orde komen (tijdstip van thuiskomst, mee op vakantie gaan, welke tv-programma's kijken). De vraag: 'Jongens en meisjes van 16 jaar mogen zelf bepalen hoe laat ze thuis komen', is gesteld aan jongeren tussen 13 en 19 jaar.

De resultaten van beide landelijke onderzoeken en van ons eigen onderzoek met betrekking tot zeggingskracht over het tijdstip van thuiskomen zijn weergegeven in Tabel 5.

Hieruit blijkt een duidelijke discrepantie tussen de door de jongeren zelf geclaimde autonomie en de bereidheid van de moeders deze autonomie te verlenen, waarbij opvalt

Tabel 5 *Autonomieverlening m.b.t. het tijdstip van thuiskomen (percentages)*

	EEN ZESTIENJARIG KIND			EEN TWINTIGJARIGE DOCHTER	
Arkel/Kedichem 1986: Gehuwde moeders	Landelijk: NOJ, 1987 13-19 jarigen			Landelijk: SCP, 1986 16-24 jarigen	
Wij laten het kind:	Kind mag het tijdstip zelf bepalen:			Bepaling van het tijdstip:	
Helemaal vrij 5		jongens	meisjes	aan dochter overlaten	43
Behoorlijk vrij 30	(helemaal) mee eens	30	22	samen	48
Enige vrijheid 45	niet eens/niet oneens	37	34	ouders	9
Niet meebeslissen 20	(helemaal) niet mee eens	33	44		

dat jongens die claim sterker leggen dan meisjes. Meisjes zijn over het algemeen meer coöperatief en conformerend dan jongens (vgl. Peterson & Rollins, 1987, p. 485). Met de nodige reserves vanwege verschillen in vraagstelling, kunnen we eveneens concluderen voor zaken als de keuze van televisieprogramma's en vakantiebestemming dat moeders zeggen een grotere vrijheid van keuze te geven dan de kinderen zelf zeggen te krijgen.

Terugkerend naar onze probleemstelling kunnen we stellen dat ouders er een gedifferentieerd patroon van autonomieverlening op na houden. Ouders blijken bij de meeste zaken scherp naar leeftijd te differentiëren. De AUTONOMIEVERLENING daalt bij stijgende leeftijdscategorie; maar die daling is tevens afhankelijk van de aard van het vrijetijdsdomein. De keuze van een sport is daarbij het terrein waarop ouders hun kinderen het meest vrij laten. Er is geen enkele ouder die hierover volledige zeggenschap claimt wanneer het kind 12 of 16 jaar is. Dit in tegenstelling tot het tijdstip van thuiskomen. Er is geen enkele ouder die een kind van 8 of 12 jaar het tijdstip van thuiskomen volledig zelf laat bepalen.

Met betrekking tot het tijdstip van naar bed gaan en het al dan niet mee op vakantie gaan, handelen ouders relatief strikt, inzake het logeren bij een vriend(in), de vraag wie mag komen logeren, de keuze van TV-programma's en wat te doen tijdens de vakantie, handelen ze over het algemeen wat minder strikt.

De analyse laat verder zien dat er samenhang is waar te nemen in de houding ten aanzien van het groeien naar zelfstandigheid. Naarmate ouders op het ene terrein meer resp. minder invloed laten gelden, doen ze dat ook op andere terreinen. De 27 door ons geformuleerde items leverden een zinvol te hanteren meetinstrument inzake AUTONOMIEVERLENING op. Het meten van AUTONOMIEVERLENING voor verschillende leeftijden van de kinderen is daarbij een belangrijke versterking van de schaal gebleken. Tevens is gebleken dat er twee deelpatronen van opvoedingsgedrag inzake zelfstandigheid van kinderen zijn te ontdekken (zie Figuur 5).

Uit de literatuur is bekend dat ouders vrije tijd van zichzelf en van hun kinderen niet (in de eerste plaats) alleen zien als tijd voor recreatie en zelfontplooiing, maar ook – en juist uitdrukkelijk – als pedagogische ruimte. In en door middel van vrije tijd proberen ouders pedagogische doeleinden te verwezenlijken. Ouderlijke invloed op beslissingen met betrekking tot die vrije tijd wordt hiervoor ingezet. Interessant is dan dat ouders bij de mate van invloed onderscheid maken tussen 'invloed op zaak' versus 'invloed op tijd'. Wellicht mogen we de variatie in ouderlijke invloed op de activiteiten in verband brengen met wat de ouders zelf belangrijk vinden dat kinderen in de vrije tijd doen, terwijl 'invloed op tijd' meer te maken lijkt te hebben met het ouderlijke probleem van controle en orde, met name in de betekenis van tijdsorde. Ouders percipiëren dit aspect van invloed als meer of minder belangrijk naarmate zij meer of minder belang hechten aan ouderlijke controle en ordelijke tijdsstructuur überhaupt. Immers, een groot gedeelte van het gezinsleven wordt gestructureerd via een ordelijke

	AUTONOMIEVERLENING m.b.t tot vrije tijd als activiteit	AUTONOMIEVERLENING m.b.t. vrije tijd als tijd
1. Inhoud van de zeggenschap	Zeggenschap over activiteiten	Zeggenschap over tijd
2. Specificiteit	Meer onderwerpsspecifiek	Meer leeftijdsspecifiek
3. Aard van de zeggenschap	Vraag òf een kind iets mag	Vraag wannéér een kind iets mag
4. Mate van autonomieverlening	Gemiddeld en relatief laag	Gemiddeld en relatief hoog
5. Controlemogelijkheid van de ouders	Relatief groot	Relatief gering.

Figuur 5 *Twee deelpatronen van opvoedingsgedrag inzake zelfstandigheid van kinderen*

tijdsindeling, waarbij de 'agenda's' van moeder, vader en kinderen synchroon dienen te lopen.

Het probleem met betrekking tot de vraag welke factoren van invloed zijn op de mate AUTONOMIEVERLENING heeft zich vooral gericht op het belang van sociale klasse en sekse. Kiezen we als indicator voor sekse de sekseverdeling van kinderen in de onderzochte gezinnen, dan blijkt deze sekseverdeling aan de verklaring van de variatie in AUTONOMIEVERLENING meer bij te dragen dan sociale klasse. De schaal voor AUTONOMIEVERLENING blijkt in feite onafhankelijk van de sociaal economische status van een gezin te zijn. Nadere analyse laat zien dat via 'bemiddeling' van de opvoedingswaarden ZELFBEPALING en SEKSE-ONDERSCHEID de AUTONOMIEVERLENING indirect gerelateerd kan worden aan de milieuvariabelen sociaal economische status en geloofsparticipatie, terwijl de sekseverdeling van de kinderen een zelfstandige factor blijft. Onze bevindingen sluiten daarmee aan op hetgeen Gecas (1979), Hynes (1985) en ook Gerris en Janssens (1987) concluderen ten aanzien van het model van Kohn. We trekken hieruit de conclusie dat sekse – in verschillend opzicht – voor de verklaring van verschillen in opvoedingsgedrag (wat betreft de toe te kennen beslissingsvrijheid op vrijetijdsdomeinen) belangrijker is dan sociaal milieu; dit strookt ook met de bevindingen van Kamphorst en Spruijt (1983) daaromtrent.

Een meer definitieve conclusie zou een onderzoeksontwerp toelaten, waarin naast de moeder ook de vader alsmede jongens èn meisjes, uit verschillende sociale klassen onderwerp worden van een soortgelijke studie als de onze. Dan kunnen ook een meer directe vergelijking en weging gemaakt worden tussen door de vader/moeder te verlenen autonomie en de door jongens/meisjes geclaimde autonomie op onderscheiden vrijetijdsdomeinen.

Literatuur

Admedia B.V. & Oberon B.V., *Nationaal onderzoek jongerenbladen*. Amstelveen: VNU Marktonderzoek, 1987.

Bosker, R.J., G.W. Meijnen & R.K.W. van der Velden, *Waardenoriëntaties en maatschappelijke gelaagdheid*. Paper voor de studiedagen voor sociologen en antropologen. Amsterdam, 1986.

Gecas, V., The influence of social class on socialization. In: W.R. Burr, R. Hill, F.J. Nye & J.L. Reiss (Eds.), *Contemporary theories about the family, Vol. 1*. London: Free Press, 1979.

Douma, W.H., *Het gezin tussen verleden en toekomst. Een sociologisch onderzoek in twee verstedelijkte plattelandsgemeenten voor en na de jaren zestig: beschrijving, analyse en beschouwingen*. Wageningen: Veenman & Zonen, 1975.

Gerris, J.R.M. & J.M.A.M. Janssens, Opvoedingsgedrag van ouders in overtredingssituaties: een onderzoek naar beïnvloedende factoren. *Pedagogische Studiën*, 1987, *64*, 7/8, 295-308.

Heiden, A. van der & R. Smeding, *Dagrekreatie en het gezin*. Amsterdam: SISWO, z.j.

Hynes, E., Socialization values and punishment behavior; An Irish test of Kohn's thesis. *Sociological Perspectives*, 1985, Vol. 28, no. 2, 217-239.

Jansen, H., Klasse- en sexespecificiteit in de waardering van conformiteit en zelfbepaling; enkele ervaringen met de schaal van Kohn. *Sociologische Gids*, 1983, 1, 37-44.

Jöreskog, K.G. & D. Sörbom, *Lisrel VI; Analysis of Linear Structural relationships by the method of maximum likelyhood, Instrumental Variables, and Least Squares Methods*. Mooresville: Scientific Software Inc., 1985.

Kamphorst, T. & E. Spruijt, *Vrijetijdsgedrag in het perspectief van socialisatie*. Utrecht: Rijksuniversiteit Utrecht, 1983.

Knijn, T. & C. Verheijen, *Moederschapsopvattingen, dagelijkse praktijk en spanningen*. Nijmegen: Katholieke Universiteit Nijmegen, 1986.

Kooy, G.A., *Sexualiteit, huwelijk en gezin in Nederland*. Deventer: Van Loghum Slaterus, 1975.

Maccoby, E.E. & J.A. Martin, Socialization in the context of the family: Parent-child interaction. In: P.H. Mussen (Ed.), *Handbook of child psychology, Vol. 4*. New York: John Wiley & Sons, 1983.

Meijnen, G.W., *Maatschappelijke achtergronden van intellektuele ontwikkeling*. Groningen: Wolters Noordhoff, 1979.

Meijnen, G.W. *Van zes tot twaalf; een longitudinaal onderzoek naar de milieu- en schooleffecten van loopbanen in het lager onderwijs*. 's-Gravenhage: SVO reeks no. 79, 1984.

Newson, J. & E. Newson, *Seven years old in the home environment*. London: George Allen & Unwin, 1976.

Peterson, G.W. & B.C. Rollins, Parent – child socialization. In: M.B. Sussman & S.K. Steinmetz (Eds.), *Handbook of marriage and the family*. New York: Plenum, 1987.

Rollins, B.C. & D.L. Thomas, Parental support, power and control techniques in the socialization of children. In: W.R. Burr, R. Hill, F.J. Nye & J.L. Reiss (Eds.), *Contemporary theories about the family, Vol. 1.* London: Free Press, 1979.

Sociaal en Cultureel Planbureau, *Sociaal en Cultureel rapport 1986.* 's-Gravenhage: Staatsuitgeverij, 1986.

Swanborn, P., *Schaaltechnieken. Theorie en praktijk van acht eenvoudige procedures.* Meppel: Boom, 1982.

Vervoort, C.E., *Gezin en schoolkeuze bij handarbeiders.* Leiden, RUL, 1968.

Walters, J. & N. Stinnett, Parent-Child relationships: A decade review of research. *Journal of marriage and the family,* 1971, *1*, 70-111.

Wouw, B.A.J. van der, *Huishouden en ouderschap.* (nog te publiceren).

Westerlaak, J.M., J.A. Kropman & J.W.M. Collaris, *Beroepenklapper.* Nijmegen: ITS, 1975.

Curricula vitae

B.A.J. van der Wouw (1955) studeerde sociologie aan de Landbouwuniversiteit Wageningen. Werkzaam geweest als universitair docent bij de vakgroepen Sociologie van de westerse gebieden en Huishoudkunde van de Landbouwuniversiteit Wageningen. Sinds 1988 als sociaal wetenschappelijk onderzoeker verbonden aan het Provinciaal Opbouworgaan Stichting Zeeland. Verricht promotieonderzoek naar de relatie tussen huishouden en ouderschap.

J.W. te Kloeze (1945) studeerde sociologie aan de Katholieke Universiteit Nijmegen, specialisatie Sociologie van gezin en sociale welzijnszorg. Is universitair docent (recreatiesociologie en methodenleer) bij de vakgroep Sociologie van de westerse gebieden van de Landbouwuniversiteit Wageningen. Onderzoek en publikaties hebben betrekking op gezin en vrije tijd en op kleinschalig kamperen in Nederland.

Adres: Vakgroep Sociologie van de Westerse Gebieden, Landbouwuniversiteit Wageningen, Hollandseweg 1, 6706 KN Wageningen

Manuscript aanvaard 29-1-'90

Summary

Wouw, B.A.J. van der, J.W. te Kloeze. 'Children, autonomy and leisure'. *Pedagogische Studiën*, 1990, *67*, 368-381.

In our society leisure mostly takes place in domestic domains or with the family. For structuring their leisure children are largely dependent on their parents' consent and approval. In this article a measuring instrument is presented which indicates the autonomy parents give regarding their children's leisure.
This instrument is used to test Kohn's theoretical model explaining differences in parenting styles. In contrast with Kohn's thesis, in this study we found no direct relationship between social class and given autonomy. Strong relationships were found, however, between social class and parental values and between parental values and given autonomy. Also family structure is found to have some influence on given autonomy. Mothers differentiate between control of leisure time (when a child is allowed to do something) and control of leisure activities (whether a child is allowed to do something). They are more strict with respect to leisure time.

Boekbespreking

W. Koops & J. J. van der Werff (red.), *Overzicht van de empirische ontwikkelingspsychologie, deel 3: Sociale ontwikkeling en de ontwikkeling van de persoonlijkheid.* Wolters-Noordhoff B.V., Groningen, 1988, 383 pag. *f* 69,50, ISBN 00-01-48898-6.

In 1979 verscheen het overzicht van de ontwikkelingspsychologie onder redactie van W. Koops en J. J. van der Werff. Zo'n tien jaar later rondden beide redacteuren de herziening van dit overzicht van hun vakgebied af met een deel dat de sociale en persoonlijkheidsontwikkeling tot onderwerp heeft. Hieraan vooraf gingen twee andere delen. Het eerste deel, uitgegeven in 1987, behandelt de grondkenmerken van het vak. Het tweede deel, ter perse in 1988, gaat over de ontwikkeling van functies en cognitie.

Van één deel naar drie delen dus om de empirische ontwikkelingspsychologie overzichtelijk te kunnen neerzetten. Kan nu ook, met een dergelijke groei in betrekkelijk korte tijd, gesproken worden van een veelbelovende ontwikkeling van het onderhavige vakgebied? Bij de beantwoording van deze vraag zal ik me beperken tot de inhoud van deel 3. Eerst geef ik beknopt aan welke onderwerpen in dit deel de revue passeren. Vervolgens zal ik kort uiteenzetten welke plaats deze uitgave van het Overzicht naar mijn mening verdient.

Strikt genomen omvat het onderzoek naar de sociale en persoonlijkheidsontwikkeling een bonte verzameling van uitkomsten die verwijzen naar de ontwikkeling van de mens als uniek individu en tegelijkertijd ook sociaal wezen. Een overzichtelijke presentatie van zo'n verzameling van gegevens vereist enige structurering. Koops en Van der Werff kiezen hiertoe, in het eerste hoofdstuk van het boek, vooral voor een historische ordening. Betoogd wordt dat het onderzoek naar de sociale ontwikkeling en de ontwikkeling van de persoonlijkheid aanvankelijk vooral beïnvloed is door ideeën uit de psychopathologie (in hoofdzaak Freud en zijn identificatietheorie) en het behaviorisme (in hoofdzaak de sociale leertheorie van Bandura). Meer recent is vooral de nadruk op cognitieve en biologische aspecten en determinanten van de sociale en persoonlijkheidsontwikkeling merkbaar.

Een andere manier om het onderzoek naar de sociale en persoonlijkheidsontwikkeling te ordenen, is door het onderzoek te plaatsen op de nature-nurture dimensie: ligt de nadruk op genetische of omgevingsinvloeden. Het is deze dimensie die in belangrijke mate ten grondslag ligt aan de volgorde van onderwerpen die in het boek aan de orde komen. Biologische en fysiologische gezichtspunten voeren de boventoon in het hoofdstuk 2 (biopsychologie van de sociale ontwikkeling) hoofdstuk 3 (temperament als basis van de sociale ontwikkeling) en, zij het afnemend, in hoofdstuk 4 (de ontwikkeling van gehechtheidsrelaties). Omgevingsinvloeden staan centraal in hoofdstuk 8 (de televisie als socialiserende factor) en het afsluitende hoofdstuk 9 (sociale ontwikkeling). De drie nog niet genoemde hoofdstukken zijn vooral illustratief voor de toenemende aandacht voor sociale aspecten van de cognitieve ontwikkeling (hoofdstuk 5: ontwikkeling van sociale cognitie) en voor de cognitieve aspecten van de sociale ontwikkeling (hoofdstuk 6: zelfconceptie in ontwikkelingspsychologisch perspectief en hoofdstuk 7: de ontwikkeling van het morele denken). De beide ordeningen (historisch en ééndimensionaal) hebben als functie dat de onderwerpen die in het boek besproken worden niet louter opgevat worden als een betrekkelijk willekeurig gekozen verzameling van empirische bevindingen die tot het terrein van de sociale en persoonlijkheidsontwikkeling gerekend kunnen worden. Dit laat evenwel onverlet dat verschillende onderwerpen in het boek niet, beperkt, of ingekaderd binnen een andere thematiek aan de orde komen. Zo wordt bijvoorbeeld de ontwikkeling van agressief en pro-sociaal gedrag besproken in de context van de televisie als socialiserende factor (hoofdstuk 8). In die context is de uitspraak dat er per traditie relatief weinig onderzoek is verricht naar pro-sociaal gedrag te billijken (pag. 247). In zijn algemeenheid doet deze uitspraak evenwel te weinig recht aan hetgeen hierover ook vanuit een empirisch ontwikkelingspsychologisch perspectief te vermelden zou zijn. Een ander voorbeeld betreft de emotionele ontwikkeling, die wat versnipperd terug te vinden is bij emotionele sociale cognitie (pag. 130-132), de ontwikkeling van gehechtheidsrelaties (hoofdstuk 4), temperamentstrekken (hoofdstuk 3) en in zekere zin ook bij vriendschapsrelaties (hoofdstuk 9, pag. 275-279). Geen aandacht krijgt bijvoorbeeld de ontwikkeling van het prestatiemotief. Een reden hiervoor kan zijn dat in het overzicht geen ruimte is gereserveerd voor de invloed van de school als socialiserende factor. Een tipje van de sluier op dit terrein wordt opgelicht in hoofdstuk 9 als er gesproken wordt over verschillen in populariteit (pag. 265-267).

Ook over de volgorde waarin de onderwerpen aan de orde komen, is een opmerking te maken. Vanuit didactisch oogpunt bezien, verdient het aanbeveling te starten met een plaatsbepaling van de sociale en persoonlijkheidsontwikkeling binnen de

empirische ontwikkelingspsychologie. De redacteuren slagen hierin voor wat betreft de historische gang van zaken. Ze blijven echter in gebreke het terrein ook meer begripsmatig te preciseren. Wat rekenen we tot de sociale ontwikkeling? En wat is kenmerkend om van de ontwikkeling van de persoonlijkheid te kunnen spreken? Welke begrippen en theorieën spelen daarbij ook heden ten dage nog een belangrijke rol? En welke verschijnselen kunnen vanuit deze theoretische kaders beschreven en verklaard worden? De inleiding van het laatste hoofdstuk (sociale ontwikkeling) geeft voor een deel weer wat ik meer in het begin zou wensen. Het begrip sociale ontwikkeling wordt omschreven en in perspectief geplaatst van de ontwikkeling van de eigen individualiteit of, zo men wil, de persoonlijkheid.

Ik sluit af. Het boek bevat onderwerpen die stuk voor stuk de moeite waard zijn. Een lezer krijgt een redelijk overzicht, op inleidend niveau, van de stand van zaken per onderwerp. Als zodanig verdient het boek een plaats in de boekenkasten en het denken van degenen die besluiten zich te verdiepen in de ontwikkelingspsychologie. Daar staat tegenover dat de onderwerpen die besproken worden maar zeer gedeeltelijk het gebied bestrijken dat tot de sociale en persoonlijkheidsontwikkeling gerekend kan worden. En daarmee is het boek onvoldoende representatief voor de ontwikkelingen op onderhavig deelterrein van de ontwikkelingspsychologie.

Hiermee is niet gezegd dat voor een overzicht moet gelden dat een terrein volledig bestreken wordt, er nog even van afgezien of dit te realiseren zou zijn binnen een handzaam formaat. Wel dat mijns inziens, meer nadrukkelijk dan in het onderhavige boek is gebeurd, aangegeven moet worden op welke gronden (historisch, maatschappelijk, begripsmatig, pragmatisch) de keuze voor de onderwerpen is gebaseerd. Het komt mij voor dat een aanzet hiertoe vereist is om dit overzicht van de sociale ontwikkeling en de ontwikkeling van de persoonlijkheid ook een plaats te laten innemen in curricula en het denken van collegae psychologen en pedagogen.

J. L. van der Linden

J. C. Voogt, *Scholen doorgelicht. Een studie over schooldiagnose,* Rijksuniversiteit Utrecht, Utrecht, 1989. ISBN 90-720-1520-7

In het proefschrift in kwestie wordt een plaatsbepaling gegeven van zogenoemde 'schooldiagnose-modellen' binnen de literatuur over onderwijsinnovatie, schoolorganisatie en evaluatie. Schooldiagnose-modellen behoren tot de familie van benaderingen die ik zelf liever met de term 'instellingsspecifieke evaluatie' zou aanduiden. Binnen de verschillende benaderingen en technieken die er op dit terrein bestaan wordt wel onderscheid gemaakt tussen op schoolverbetering gerichte aanpakken en procedures die worden toegepast om tegemoet te komen aan externe controle. De benadering die in dit proefschrift centraal staat kiest uitdrukkelijk voor de eerste benadering. De term zelfdiagnose is een bewuste keuze om uit te drukken dat beoordelen hier geheel in het teken staat van het verbeteren van de school. Naast het conceptuele deel bevat het proefschrift de beschrijving van de ontwikkeling van een schooldiagnose-instrument, getiteld "Systematische Analyse ten behoeve van Schoolverbeteringsprocessen" (SAS) en van onderzoek naar de interne consistentie en de validiteit van het instrument.

Als geheel is het proefschrift te kwalificeren als relevant en 'up to date' als het gaat om de selectie van de gebruikte literatuur, netjes wat betreft de uitvoering van het valideringsonderzoek en helder en prettig lezend wat de presentatie betreft. De verdienste van het verrichte ontwikkelingswerk is dat er een voor de praktijk bruikbaar instrument voor zelfonderzoek tot stand is gebracht, terwijl het instrument ook gebruikt kan worden voor interessante mogelijkheden tot vervolgonderzoek (zoals onderzoek naar de vraag of scholen die zichzelf een hoge score geven op het instrument ook beter presteren en onderzoek naar de vraag of scholen die het SAS-instrument gebruikt hebben ook daadwerkelijk verbeteren).

Desalniettemin zijn er ook enkele meer kritische kanttekeningen te maken. De taxonomie van schooldiagnose-modellen, waarbij onderscheid gemaakt wordt tussen een traditioneel, een democratisch, een liberaal en een kritisch model, doet op onderdelen wat geforceerd aan, in die zin dat de definiërende kenmerken lang niet altijd als onderling uitsluitend kunnen worden gezien. Zo zou, naar mijn mening, bijvoorbeeld beslissingsgerichte evaluatie evengoed met het 'traditionele' model kunnen worden geassocieerd als doelgerichte evaluatie.

Een ander onderdeel van het conceptuele deel van het proefschrift waar ik moeite mee heb, is de stelling dat er meerdere gelijkwaardige 'soorten' kwaliteit van organisaties zijn. Zo worden naast 'output-kwaliteit' input- en proces-kwaliteit onderscheiden. Naar mijn mening is de kwaliteit van organisatiestructuur en technologie binnen een organisatie afhankelijk van de vraag of daarmee op doeltreffende en doelmatige wijze wordt bijgedragen aan de primaire opbrengsten van de organisatie (in scholen zijn dat leerresultaten). Proces-kwaliteit als een op zich staand normatief criterium beschouwen, leidt tot denken in de trant van 'operatie

geslaagd, patiënt overleden'. De 'proces-kwaliteit' die met het SAS-instrument vastgesteld wordt zal zich nog moeten bewijzen, in die zin dat nog empirisch vastgesteld moet worden of scholen die hoog scoren op dit instrument inderdaad ook betere leerresultaten weten te realiseren.

Niet zozeer een punt van kritiek als wel een punt dat de verbazing opwekt, is de uitkomst van het betrouwbaarheidsonderzoek dat vijf van de zeven subschalen van het instrument uiteindelijk een redelijke interne consistentie bleken te hebben. Wanneer men kijkt naar de nogal heterogeen aandoende verzameling van kenmerken binnen de factoren (pp. 129-131) is dit opmerkelijk.

Ten slotte moet men zich goed realiseren dat het SAS-instrument een diagnose geeft die louter op opinies is gebaseerd. Naar mijn mening is het wenselijk dat instrumenten voor instellingsspecifieke evaluatie tevens objectieve gegevensverzameling behelzen. Ook dan blijven opinies van docenten relevant, maar ze hebben dan een duidelijker feitelijke basis.

J. Scheerens

Vera Asselbergs-Neessen, *Kind, kunst en opvoeding: De Nederlandse beweging voor beeldende expressie.* Acco, Amersfoort/Leuven, 1989, 317 pag., *f* 49,-, ISBN 90 5256 004 8.

Het tekenlokaal van de progressieve school uit de vijftiger jaren was gul toegerust met grote vellen papier, forse potten plakkaatverf en brede kwasten. Op een schoolbord stonden fantasie-prikkelende onderwerpen geschreven, waaruit men kon, maar niet hoefde kiezen, zoals 'droombos', 'circuspaard' of 'storm aan zee'. Wij leerlingen konden het moeilijk te bont maken: zolang ons werk er vrij uitzag, was het prachtig. Van hier uit verhuisde ik naar een minder vooruitstrevende school: de papiermaat was nu tot een kwart gereduceerd en wij werkten met fijne penselen. Voor wij begonnen met tekenen moest er eerst rondom een rand van 1 centimeter met potlood en lineaal afgetekend worden; rechtsonder diende onze naam te verschijnen in blokletters van 3 millimeter hoog. Daarna tekenden wij een voorbeeld van het bord na, of we maakten kleurstalen, een gang met zwart-witte vloertegels in perspectief of zowaar een ontwerp voor een gordijnstof. Toen ik anderhalf jaar later weer naar de progressieve school terugkeerde, wachtten daar opnieuw die grote bladen, blijmoedig opbollend en krinkelend onder hun overmaat aan plakkaatverf.

Zowel vanwege mijn herinneringen aan alle vaak tegenstrijdige kunstpedagogische methodes waar ik in mijn jeugd door beïnvloed werd, als vanwege mijn huidige belangstelling in kunsteducatie, nam ik met meer dan een gewone belangstelling "Kind, kunst en opvoeding" van Vera Asselbergs-Neessen ter hand. Ik denk dat degene die een kind op het kunstzinnige pad wil leiden over een aantal mogelijkheden en inzichten moet beschikken: over vakkennis allereerst, over inzicht in de kinderlijke ontwikkeling bovendien en over een leidende filosofie tot slot. Het is met dit laatste aspect van de gedachten en idealen over kunst en opvoeding dat Asselbergs zich bezighoudt.

Haar werk is gericht op het ontstaan en de ontwikkeling van de Nederlandse expressie-beweging tussen 1876 en 1968 in een internationaal perspectief. Het centrale thema van de vrije expressiebeweging ziet zij als de spanning tussen vrijheid van expressie en ambachtelijke gebondenheid, tussen culturele groei en cultuuroverdracht.

Het boek is in drie delen opgebouwd. In het eerste deel behandelt Asselbergs de periode 1876-1918, waarin een ontluikende belangstelling optreedt voor de kinderlijke expressie. In deze periode werd het werk van beeldende kunstenaars opgevat als een uitdrukking van hun eigen identiteit, meer dan als een afgeleide van de optische werkelijkheid. Gauguin bijvoorbeeld was op zoek naar datgene wat nog niet door de Westerse burgerlijke samenleving was aangetast: de kinderlijke oorspronkelijkheid, de ongerepte natuur en de exotische culturen. Een reeks opvoedkundige stromingen begon nu een pedagogische primitivisme te prediken, waarbij het pure, onschuldige kind het ideaalbeeld werd voor de 'nieuwe mens'. Vanaf 1885 verschenen er ook verschillende studies over het eigene in de kindertekening, o.a. van de Engelsen Cooke en Sully, van de Italiaan Ricci en van de Duitser Kerchensteiner. Vanaf 1892 raakte de Oostenrijkse pionier Franz Cizek zeer geïnteresseerd in het 'natuurlijke tekenen'. In zijn beroemde "Jugendkunstklasse" in Wenen benadrukte hij vooral *vrijheid*, een idee dat in Nederland voorlopig nog maar een zwakke weerklank vond.

In de volgende periode die Asselbergs onderscheidt, van 1919-1945, werd internationaal en nationaal gestreden voor het behoud en de ontwikkeling van de vrije expressie van kinderen. In Nederland werden de ideeën van Kerchensteiner, Montessori, Steiner e.a. bekend. Vanuit de kunstwereld waren het surrealisme, het expressionisme en de opvattingen van het Bauhaus inspiratiebronnen. Kees Boeke werd nu een leidende figuur van de pedagogische avant-garde met zijn "Werkplaats Kindergemeenschap"; Jan Jordens, tekenleraar en lid van het kunstenaarsgenootschap "de Ploeg", gaf zijn leerlingen veel vrijheid om de ontwikkeling van hun spontane expressiviteit niet te remmen.

In de kindertekening werd nu ook het psychologische aspect van de vrije expressie ondekt. Terwijl

B. Merema het onderbewuste in kindertekeningen wilde onderzoeken, de 'donkere broeder', kwamen tekenleraren in Nederlands Indië onder de indruk van de artistieke produktie van de 'kinderen van het oerwoud'(!). In het buitenland werden de publikaties invloedrijk van Herbert Read en van Victor Löwenfeld, die in Wenen met Cizeks Jugendkunstklasse kennis had gemaakt.

De derde periode, van 1945-1968, karakteriseert Asselbergs als de doorbraak en verwatering van de vrije expressie-beweging. In de periode direct na de oorlog ging er een stimulerende rol uit van het Stedelijk Museum in Amsterdam en speciaal van de persoon van Willem Sandberg, die er in 1948/49 een tentoonstelling "Kunst en Kind" organiseerde. De kunstenaars van Cobra lieten zich in deze jaren door de kindertekening inspireren. De kortstondige vreugde van de bevrijding (die spoedig overging in een ascetische moraal van aanpakken en vooruitzien, nodig voor de wederopbouw) zorgde voor een doorbraak in de vrije expressiebeweging, die sterk naar voren kwam in de Nutsvolksschool voor beeldende kunsten en de Werkschuit.

Hoewel ieder individu en iedere groep eigen accenten legde, hadden de uitgangspunten voor het stimuleren van de vrije expressie overal grote overeenkomst. *Vertrouwen* in de scheppende kracht van het kind, een discrete, op de achtergrond blijvende *behulpzaamheid* bij techniek en materiaalgebruik, *prikkelen* van de fantasie, de waarneming en de herinnering en *bevorderen* van de *eigen* uitdrukkingwijze van de *eigen* ervaring waren sleutelbegrippen. Een laissez-faire didactiek werd algemeen afgewezen: men streefde doelbewust naar de bevrijding van beeldende cliché's uit onderwijs en massacultuur.

In het laatste hoofdstuk van haar boek plaatst Asselbergs alle personen, bewegingen en instellingen die van belang zijn geweest voor de vrije expressie-beweging in een overzichtelijk schema. Uiteindelijk schetst zij de vernieuwing in het kunstonderwijs als een beweging in een vlak bepaald door twee haaks op elkaar staande dimensies: die van cultuuroverdracht versus culturele groei en die van collectiviteit versus individualiteit. De vernieuwingsbeweging van het expressie-onderwijs verloopt hierin nu 'diagonaalsgewijze' van collectieve cultuuroverdracht naar individuele groei.

Het uiteindelijke verwateren van de vrije expressiebeweging ziet Asselbergs als voortkomend uit een tekort aan beeldend vakmanschap en een steeds meer verwaarlozen van pedagogiek en didactiek in de creativiteitscentra. Er ontstonden nieuwe cliché's: was vroeger de 'gemberpot met judaspenningen' een verplichte opgave, nu was 'de gekke kop' of 'de wonderbloem' een standaard geworden. Voor de toekomst is het volgens Asselbergs dan ook nodig meer eisen te stellen aan *kwaliteit*: zowel aan die van de docenten op alle niveaus, als aan die van de werkstukken die onder hun begeleiding ontstaan.

Vera Asselbergs heeft een zeer gedegen en gedetailleerde kunst-sociologische studie gemaakt over alle personen, groepen en groepjes, stromingen en tegenbewegingen, die bij de vrije expressiebeweging een rol hebben gespeeld. Deze plaatst zij tegen een cultuurhistorische achtergrond, bij de lezer een besef oproepend hoezeer onderwijs-idealen geworteld zijn in een bredere maatschappelijke context. Voor de wat minder op detail ingestelde lezer zijn bij alle hoofdstukken en delen samenvattingen gegeven, waarmee het bomen-en-bos probleem doeltreffend is opgelost.

De vormgeving van het boek met talrijke illustraties in zwart-wit en kleur is bijzonder aantrekkelijk. Manifesten verouderen, maar kindertekeningen niet – zelfs niet als ze gemaakt zijn volgens inmiddels vergeelde principes. Terwijl de kleurige tekeningen van de Werkschuit ons opvrolijken, hebben die foto's van meisjes in lange boezeroenen die golflijnen op het schoolbord tekenen toch iets nostalgisch-vertederends, al ziet men hun krijtlijnen nu als uit de tijd en uit den boze. Vera Asselbergs' boek verdient het een standaardwerk te worden voor al diegenen die zich met de expressievakken bezig houden.

A.Th. van Meel-Jansen

Graaff, E. de, *Simulation of Initial Medical Problem-solving*. Uitgeverij Thesis, Amsterdam, 1989, 86 pag., *f* 35,-. ISBN 90 5170 006 7.

Zelden begon een promovendus zijn proefschrift met een zo toepasselijk citaat als De Graaff: "No problem is so big or complicated that it can't be run away from. (Charley Brown)". De auteur is er namelijk in geslaagd het schrijven van een dissertatie te vermijden en toch te promoveren.

De beoordeling van de kwaliteit van een proefschrift is natuurlijk een tamelijk subjectieve zaak en blijkbaar verschil ik hier van mening met de promotor R. Drop, hoogleraar Medische Sociologie aan de R.U. Limburg.

Het proefschrift heeft gelukkig een minimale omvang, 86 bladzijden met veel wit. Zowel voor een goed als een slecht boek is een geringe omvang een voordeel. Het siert een auteur die erin slaagt in een beperkte ruimte zijn orginele ideeën en scherpe analyses helder te presenteren. Een slecht boek kan natuurlijk niet kort genoeg zijn. In deze 86 bladzijden is het bijna onmogelijk om een aardig idee of een minstens correct gepresenteerde en geïnterpreteerde analyse te vinden.

De verantwoording van een toets, SIMP ge-

naamd (zie titel), is het hoofdonderwerp van het proefschrift. Medische studenten wordt daarin hun eerste reactie als huisarts gevraagd op een tiental casusbeschrijvingen van ongeveer 15 regels. De reactie wordt schriftelijk gegeven. Omdat geen werkelijke patiënt-ontmoetingen plaatsvinden wordt 'Simulatie' aan de naam van de toets toegevoegd. De kwalificatie 'Problem-solving' wordt gebruikt, omdat niet rechtstreeks naar medische kennis wordt gevraagd, maar om een adequate toepassing daarvan in een praktijksituatie. Er vindt geen analyse of beoordeling plaats van cognitief gedrag, noch een simulatie met een computermodel. Slechts de adequaatheid van het antwoord wordt beoordeeld. Gewone wis- en natuurkunde-toetsen voor begrijpend lezen of schrijven zouden derhalve eveneens met de attributen 'simulatie' en 'problem-solving' kunnen worden opgetuigd. Ook hier gaat het immers niet om werkelijke situaties en ook moeten er problemen worden opgelost.

De auteur gebruikt hoofdstuk 2 om aan te tonen dat "medical problem-solving" gelijk mag worden gesteld met "the ability to solve medical problems" (blz. 16). Aan dit inzicht verbindt hij de conclusie dat het geen zin heeft om aandacht te geven aan zoiets als een cognitief proces achter de actie. Bij Boshuizen (1989), onlangs gepromoveerd aan dezelfde universiteit, vinden we dat er in dit opzicht een verschil is tussen een student en een ervaren arts, en dat de vorming van adequate medische reacties, aanvankelijk, dus bij studenten en beginnende artsen, wel een bepaald redeneerproces vergt.

In hoofdstuk 3 wordt de toetsvorm verantwoord, met name waarom alleen aandacht aan de eerste reactie wordt gegeven. De reden daarvoor is kortweg dat een volledige arts-patiënt-interactie voor een meting, een toets, niet hanteerbaar is, en de eerste reactie in hoge mate bepalend is voor de kwaliteit van de overige reacties. Dit vergt natuurlijk enige uitleg: vandaar hoofdstuk 3. En passant demonstreert de auteur zijn kennis van de literatuur over probleem-oplossen met de gotspe dat deze beperkt zou zijn tot abstracte puzzles (blz. 28). De verwijzing naar het monument van Newell en Simon uit 1972 moet deze uitlating ondersteunen. Mettes en Pilot (1980), Schoenfeld (1985) en vele anderen worden door de Graaff genegeerd.

Met de eerste drie hoofdstukken hebben we een weinig zeggende en obligate inleiding op het onderzoek achter de rug. De laatste twee hoofdstukken, voor de conclusie in hoofdstuk 6, worden een klein beetje technisch. Dit beetje is de auteur echter veel te veel. Het gehannes met generaliseerbaarheids- en betrouwbaarheidscoëfficiënten is treurig.

Om betrouwbaarheid en validiteit van de toetsresultaten gunstig te beïnvloeden werden voorschriften ontwikkeld voor het scoren van de antwoorden. Deze voorschriften zijn zo opgesteld dat een beroep kan worden gedaan op de schaarse maar goedkope tijd van verpleegsters i.p.v. de dure tijd van werkzoekende artsen.

Het eerste deel van hoofdstuk 4 beschrijft het onderzoek naar de betrouwbaarheid van de oordelen van verpleegsters. Voor het kwantificeren van de generaliseerbaarheid van de toetsresultaten wordt in paragraaf 2 gebruik gemaakt van Intra-Class Correlaties zoals beschreven wordt in Shrout en Fleiss (1979). Paragraaf 3 maakt gebruik van Mellenbergh (1979). Deze wisseling van bron speelt de auteur parten als hij coëfficiënten gaat vergelijken.

Bij het onderzoek in paragraaf 2 zijn 25 vijfdejaars studenten en 6 verpleegsters betrokken. De studenten krijgen 22 casussen voorgelegd. Als maat voor de betrouwbaarheid wordt ICC (2,1) uit Shrout en Fleiss (1979) gekozen. In dit geval geeft deze coëfficiënt een schatting van de correlatie tussen twee oordelen over een reeks at random getrokken student-antwoorden op casussen, waarbij voor ieder afzonderlijk oordeel at random een verpleegster wordt getrokken. De auteur beweert evenwel dat ICC (2,1) de correlatie schat tussen "the scoring by one rater, drawn at random from a population of nurses, with the *mean* scoring by that population", (onderstreping HV) hetgeen onjuist en onvolledig is.

De waarde van correlatie wordt geschat op 0.83. De auteur vergelijkt deze 0.83 met de gebruikelijk voor toetsen vereiste betrouwbaarheid van 0.80. Dit criterium is echter alleen zinvol voor de betrouwbaarheid van de meting van de vaardigheid van studenten. In een toetsprocedure worden de studenten beoordeeld, niet de antwoorden op de vragen. Voor toepassing van dit criterium had de auteur een maat voor de betrouwbaarheid van de oordelen over de studenten moeten kiezen, bijv. de correlatie tussen twee oordelen over studenten door voor iedere student twee at random getrokken verpleegsters, waarbij twee keer het gemiddelde oordeel over dezelfde k casussen wordt genomen. Deze correlatie kan worden geschat door:

$$\frac{\sigma_s^2 + 1/k\,\sigma_{sc}^2}{\sigma_s^2 + \sigma_r^2 + \sigma_{sr}^2 + 1/k(\sigma_{sc}^2 + \sigma_{rc}^2 + \sigma_{scr,e}^2)}$$

Hiermee generaliseren we alleen nog maar over beoordelaars, niet over het toetsdomein.

Met behulp van de variantiecomponenten die in paragraaf 3 worden gegeven kunnen we ons enigszins een oordeel vormen over de orde van grootte ervan bij bijvoorbeeld 10 casussen, ongeveer de lengte van de toetsen. We krijgen dan 0.44. De bovengenoemde 0.83 wordt dus grotendeels veroorzaakt doordat de variantie in moeilijkheidsgraad tussen de casussen en de interactie daarvan met de studenten tot de ware variantie is gerekend.

Even verder wordt de 0.83 opgekrikt tot 0.97 door de oordelen van de 6 verpleegsters in de G-study te middelen. Nergens blijkt dat men overweegt om in de beoordelingsprocedure 6 ver-

pleegsters te gebruiken. Integendeel, wegens de kosten worden er een en hoogstens twee beoordelaars geadviseerd.

In paragraaf 3 worden 4 van de 22 casussen geselecteerd voor beoordeling door twee huisartsen. Door de titel van de paragraaf wordt gesuggereerd dat dit een aardige ingang is om na te gaan of de validiteit van de oordelen door de verpleegsters negatief wordt beïnvloed door hun geringere kennis. Met deze titel ben ik het helemaal eens. Helaas geeft de uitwerking van de paragraaf geen antwoord op de gestelde vraag. De auteur berekent de variantiecomponenten voor generaliseerbaarheidscoëfficiënten van oordelen waarbij at random wordt getrokken uit een pool van verpleegsters en artsen in de verhouding 6:2. Zelfs de minimaal verwachte correlatie tussen oordelen van verpleegsters en artsen wordt niet gepresenteerd.

De generaliseerbaarheidscoëfficiënt bovenaan blz 55, wordt ingeleid met: "Apart from the grand mean the factor cases represents by far the largest component of variance. This indicates that the reliability of the four-cases test cannot be very high. Indeed generalization to subject (replication of the measurement with the same cases and raters) gives the following result:"

$$\frac{\sigma_s^2}{\sigma_s^2 + 1/4\,\sigma_{sc}^2 + 1/8\,\sigma_{sr}^2 + 1/32\,\sigma_{scr.e}^2} = 0.42$$

Afgezien van het feit dat de grand mean niet bijdraagt aan de variantie, valt in de eerste plaats op dat de variantiecomponent van de casussen σ^2_c, in deze coëfficiënt helemaal niet voorkomt. Bovendien heeft de coëfficiënt geen betrekking op de door de auteur beschreven replicatie. De replicatie betreft dezelfde studenten, die allen de eerste keer dezelfde 4 casussen en dezelfde 8 beoordelaars krijgen, en de tweede keer wordt een nieuwe set van 4 casussen en 8 beoordelaars getrokken, die weer voor alle studenten dezelfde zijn. (De hoofdeffecten van beoordelaars en casussen ontbreken in de noemer van (2), dus geen nieuwe casussen en beoordelaars voor iedere student afzonderlijk, zoals in de tot nu toet gekozen coëfficiënten).

Deze beoordelingsprocedure (met 4 casussen en 8 beoordelaars) heeft weer geen enkele relatie tot de nagestreefde beoordelingspraktijk. Het heeft dan ook geen zin de waarde ervan te beoordelen. Toch vindt de auteur het resultaat enigszins teleurstellend, en hij neemt dan maar een andere coëfficiënt. De herhaling waarop deze coëfficiënt betrekking heeft wordt niet beschreven. Hij volstaat met: " The reliability of the raters can also be estimated from" en dan volgt de formule. Deze blijkt van dezelfde soort te zijn als de coëfficiënt die in de vorige paragraaf 0.83 opleverde. Hier wordt dat 0.80. De auteur vergelijkt hem echter met de coëfficiënt over zes beoordelaars die 0.97 opleverde en put zich vervolgens uit in het vinden van oorzaken voor het verschil tussen 0.80 en 0.97. Het onderscheid tussen zes en een beoordelaar komt daarin niet voor.

Ik bespaar de lezer dezelfde treurnis in hoofdstuk 5. Daarin worden overigens enkele interessante correlaties vermeld zoals -0.25 tussen SIMP en een medische kennistoets. De auteur concludeert hieruit dat SIMP iets toevoegt aan het meetarsenaal, omdat het iets anders meet dan medische kennis. Dat lijkt mij een juiste conclusie. Ik vind het jammer dat hij het daarbij laat. Als de correlatie juist is zou het bijvoorbeeld kunnen zijn dat de studenten die hard blokken, daarmee hun medische incompetentie trachten te compenseren. Er zullen wel meer causale hypothesen voor dit verschijnsel te verzinnen zijn. Helaas vinden we daarover niets. De reden kan niet geweest zijn dat het proefschrift te dik zou worden.

Het beperken van het testen van de vaardigheid in het stellen van een diagnose tot de eerste reactie lijkt mij een vruchtbaar idee. Dat hoeft niet verdedigd te worden met versierselen als simulatie en problem-solving. Bovendien is er aan de uitwerking daarvan in concrete toetsvragen en scoremodellen en het uitproberen daarvan door meerdere mensen veel tijd gestoken. Dit alles had een beter verslag verdiend.

H. Verstralen

Literatuur

Boshuizen, H.P.A., *De ontwikkeling van medische expertise.* Dissertatie, R.U. Limburg, 1989.

Graaff, E. de, *Simulation of Initial Medical Problemsolving.* Dissertatie, R.U. Limburg. Haarlem: Thesis, 1989.

Mellenbergh, G.J., The replicability of measures. *Psychological Bulletin*, 1977, *84*, 378-384.

Mettes, C.T.C.W. & A. Pilot, *Over het leren oplossen van natuurwetenschappelijke problemen.* Dissertatie U.T., 1980.

Schoenfeld, A.H., *Mathematical problem solving.* London: Academic Press, 1985.

Shrout, P.E. & J.L. Fleiss, Intraclass Correlations: Uses in assessing Rater Reliability. *Psychological Bulletin*, 1979, *86*, 2, 420-428.

Mededelingen

Conferentie over onderwijsverbetering

Het Instituut voor Lerarenopleiding, Onderwijsontwikkeling en Studievaardigheden (IVLOS) van de Rijksuniversiteit Utrecht organiseert op 25 januari 1991 een conferentie over onderwijsverbetering met als titel: 'Van anders naar beter. Actief inspelen op veranderingen in secundair en tertiair onderwijs'. De conferentie is bedoeld om het onderwijsveld in de gelegenheid te stellen kennis te nemen van de nieuwste inzichten over onderwijsverbetering.

Een brochure met inschrijfformulier kan worden aangevraagd bij het conferentiesecretariaat: mevr. R. Gompelman, IVLOS, Rijksuniversiteit Utrecht, Postbus 80.120, 3508 TC Utrecht, tel.: 030-532 194.

Inhoud andere tijdschriften

Comenius
10e jaargang, nr. 2, 1990

Thema: Universiteit en intellectuelen

Inleiding: een dubbelzinnige verhouding, door J. Rupp
The absent English intelligentsia, door B. Turner
Habermas over de universiteit, door M. Korthals
De idee van de universiteit-leerprocessen, door J. Habermas
Intellectuelen en politiek. Een briefwisseling, door L. Nauta en D. Pels
Derrida en de universitaire 'raison d'être', door G. Groot
De universiteit in de ogen van haar pupillen, door J. Derrida
De Gids en de nationale cultuur in de 19e eeuw, door R. Aerts
Hoger Onderwijs- en OnderzoeksPlan 1988 en 1990, door L. van Vucht Tijssen

Tijdschrift voor Onderwijsresearch
15e jaargang, nr. 3, 1990

Het oog van de meester: opstelbeoordeling door schrijvers en leerkrachten, door J.B. Hoeksma en K. de Glopper
An experimental study of the when and what of children's questions, door H. van der Meij
Het oplossen van vergelijkingsvraagstukken: een empirische toetsing van het model van Lewis en Mayer, door L. Verschaffel, E. De Corte en A. Pauwels
Notities en Commentaren
De eerste ervaringen met een geïndividualiseerde toets bestaande uit 600 vragen, door H.J.M. van Berkel, E. de Graaff, J. Sprooten en H.G. Schmidt.

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 7/8, 1990

Kanttekeningen bij de Wet op de Jeugdhulpverlening, door M. Faas
Harmonisatie van de jeugdhulpverlening, door J.D. van der Ploeg en P.M.M. Bonke
Huwelijkskwaliteit, ouderlijk handelen en (dys)functioneren van het kind, door A.J.L.L. de Brock, A.A. Vermulst en F.A.M. Leenders

Ontvangen boeken

Hertog, P.C. den, *Persoonlijkheidskenmerken en Causale Attributies* (dissertatie). Thesis Publishers, Amsterdam, 1990, *f* 49,50.
Janssen, T. & M. Overmaat, *Tekstopbouw en stelvaardigheid.* Swets & Zeitlinger, Lisse, 1990, *f* 29,50.
Petri, M., *Schoolontwikkeling. Tussen verbetering en verandering.* Acco, Leuven/Amersfoort, 1990, *f* 18,50.
Rispens, J., E. Carlier, P. Schoorl (red.), *Diagnostiek in de hulpverlening. Methodische aspecten.* Swets & Zeitlinger, Lisse, 1990, *f* 39,50.
Tisher, R.P. & M.P. Wideen (Eds.), *Research in teacher education. International perspectives.* The Falmer Press, Basingstoke, 1990, £ 12.95.
Verhoeven, L., (red.), *Etnische minderheden en geletterdheid.* Swets & Zeitlinger, Lisse, 1990, *f* 38,50.

182 TS oct 6441

Basisvaardigheden en het beroepsonderwijs*

Een operationalisering van het begrip basisvaardigheden

W. J. NIJHOF
Universiteit Twente, Enschede
J. L. M. REMMERS
IVA, Katholieke Universiteit Brabant, Tilburg

Samenvatting

Dit artikel is gericht op de vraag hoe een betere kwalitatieve aansluiting tussen onderwijs en arbeid gerealiseerd kan worden door na te gaan wat basisvaardigheden zijn; hoe deze gedefinieerd kunnen worden en eventueel operationeel kunnen worden gemaakt. De achterliggende rationale is een mogelijke taakverdeling tussen (regulier) beroepsonderwijs en andere intermediairs in de relatie onderwijssysteem-arbeidssysteem. Het vinden van kwalificaties die het karakter hebben van basisvaardigheden zouden beschouwd kunnen worden als eindtermen van een onderwijssoort of -type, waardoor een betere afgrenzing met particuliere of private opleidingen, maar ook een betere afstemming tussen onderwijssysteem en bedrijfsopleidingen mogelijk is.

1 *Het concept basisvaardigheden: een verkenning*

1.1 *Inleiding*

Uit een analyse van ontwikkelingen op de arbeidsmarkt (Nijhof & Remmers, 1989) blijkt dat functie-inhouden sterk in beweging zijn. Er treden functieverschuivingen op, zowel in de vorm van stijging als in de vorm van daling van beroepsvereisten, soms in de vorm van herschikking. Zo zorgt automatisering b.v. voor zowel een uitstoot van arbeidsplaatsen als voor het scheppen van nieuwe arbeidsplaatsen. De Amerikaanse situatie is in dit opzicht sprekend: een chronische werkeloosheid die noch met scholing, noch met 'job seeking' te bestrijden is (Wegmann, Chapman & Johnson, 1989). Deze situatie maakt het moeilijk om vooraf vast te stelllen voor welke eindtermen het onderwijs en in het bijzonder het beroepsonderwijs zou moeten voorbereiden. De behoefte aan een doorlichting en herziening van eisen ligt dan ook voor de hand.

De vraag die dan wordt geformuleerd is of er basisvaardigheden zijn die beroepsspecifiek zijn en vaardigheden die algemeen van aard zijn. Wanneer basisvaardigheden kunnen worden gedefinieerd en geoperationaliseerd voor het beroepsonderwijs, dan zou b.v. overwogen kunnen worden om de generale basisvaardigheden over te dragen via het reguliere beroepsonderwijs en de specifieke via alle mogelijke andere vormen van functiegerichte training of opleiding. Deze beleidsgerichte vraag wordt mede gestimuleerd in het licht van de kosten van het onderwijssysteem.

Uit de analyse van Nijhof en Remmers (1989) is ook af te leiden dat schoolverlaters steeds hoger opgeleid zouden moeten worden om flexibel in te kunnen spelen op de steeds veranderende eisen van de arbeidsmarkt.

De vraag is dus eigenlijk of een betere kwalitatieve aansluiting tussen onderwijs en arbeidsmarkt kan worden verkregen via basisvaardigheden. Om die vraag te kunnen beantwoorden gaan we aan de hand van literatuur na wat onder het begrip basisvaardigheden in de context van het beroep(sonderwijs) wordt verstaan, hoe ze gedefinieerd en eventueel geoperationaliseerd kunnen worden. Dan zal blijken of die vaardigheden heel algemeen, dan wel beroepsgebonden zijn.

Het zal duidelijk zijn dat we niet spreken over die vaardigheden die elke leerling leert in het algemeen vormend onderwijs, zoals lezen, schrijven en rekenen. Het gaat om vaardigheden die weliswaar een algemeen karakter hebben, maar qua niveau moeten uitstijgen boven het niveau van de algemene vorming van 16-jarigen en anderzijds voldoende raakvlakken zullen hebben met beroepssectoren waar-

* Dit artikel is gebaseerd op een studie die gefinancierd werd door het Ministerie van Onderwijs en Wetenschappen.

in b.v. technologische aspecten een rol spelen. Dat betekent dat geen enkele beroepssector of branche vooraf uitgesloten is.

Om de essentie van het begrip basisvaardigheden zo duidelijk mogelijk te formuleren is een literatuurstudie verricht. Deze sluit aan op een eerder verrichte studie in het kader van het project "Basisvaardigheden in het Beroepsonderwijs" (Nijhof & Mulder, 1986). Deze studie was bedoeld om antwoord te krijgen op de vraag welke basisvaardigheden als eindtermen zouden kunnen worden gedefinieerd en gerechtvaardigd voor het MBO (werktuigbouwkunde) en het LEAO (kantoor- en verkooppraktijk) in het licht van de moderne informatietechnologieën. De onderhavige literatuurstudie kan daarom beschouwd worden als een uitbreiding en verdere uitwerking van het begrip basisvaardigheden zoals destijds geformuleerd.

Richtsnoer voor onze analyse nu is de volgende definitie (Nijhof & Mulder, 1986): Basisvaardigheden zijn die brede kwalificerende kennis-, vaardigheids- en houdingselementen, waarover gediplomeerde schoolverlaters van het beroepsonderwijs beschikken, die transferwaarde hebben naar verwante beroepssituaties en die van essentieel belang zijn voor het uitvoeren van verschillende taken. Kernbegrippen in deze definitie zijn: de koppeling van vaardigheden aan een bepaalde doelgroep, nl. schoolverlaters van het beroepsonderwijs; voorts transferwaarde naar verwante beroepssituaties, hetgeen verwijst naar laterale of horizontale transfer. Verticale transfer wordt weliswaar niet uitgesloten, maar ook niet nader omschreven. Voorts wordt gewezen op de directe toepasbaarheid van de basisvaardigheden. Deze definitie is als werkdefinitie benut om een referentiekader op te bouwen en om een literatuursearch en -analyse te verrichten. Uit de literatuursearch kwam een veelheid aan termen en definities naar voren die aansluiten bij de gememoreerde begrippen, zoals basic skills, critical skills, generic skills, beroepsoverstijgende kwalificaties, basisvaardigheden, instrumentele vaardigheden, niet-instrumentele vaardigheden, life skills, sleutelkwalificaties, transferable skills, core skills. We zullen de gevonden literatuur en de definities systematiseren en ordenen. Het doel is te komen tot een synthese. In de volgende paragrafen zullen we achtereenvolgens een vijftal rubrieken hanteren, en wel: basic skills; job related basic skills; verruiming van het begrip basic skills; generic skills en transferable skills.

1.1.1 *Basic skills*

In de meest recente International Encyclopaedia of Educational Research (Husen & Postlethwaite, 1985) worden basic skills ook wel 'tool subjects' genoemd. Deze omvatten lezen, schrijven en rekenen. Het aanleren van deze basic skills begint al tijdens de voorschoolse periode. Een zeer belangrijk aspect van de basic skills is dat ze de basis vormen voor ieder verder leren en ontwikkelen. Er is ook een zichzelf versterkende werking bij het gebruik en het leren van basic skills. Door het lezen van een krant, een boek of studiemateriaal wordt de vaardigheid getraind en wordt kennis vergaard waarmee nieuwe informatie gemakkelijker kan worden verwerkt en geïnterpreteerd. Daarom is niet de feitelijke beheersing van de basic skills voldoende. De gebruiksmogelijkheden en het niveau van beheersing zijn buitengewoon belangrijk. Basic skills worden wel gerelateerd aan de 'back to basics movement' en ook wel aan de minimum competency testing beweging. De back to the basicsbeweging in de jaren zeventig is in de VS een reactie uit conservatieve politieke hoek op een aantal onderwijshervormingen uit de jaren zestig. Het kindgerichte onderwijs zou ineffectief zijn en wanorde en immoraliteit bij kinderen veroorzaken (Husen & Postlethwaite, 1985). De formele klassikale werkwijze op prestatiegericht onderwijs zou beter zijn. Deze overtuiging werd onderbouwd door de resultaten van zogenaamde minimum competency testing programma's die gericht waren op de drie basic skills: lezen, schrijven en rekenen. Daaruit zou blijken dat er sprake was van een algehele achteruitgang van het niveau in het onderwijs in de VS. Later is veel kritiek gekomen op de validiteit van de minimum competency tests. Een effect dat echter blijvend genoemd mag worden is een verhoging van het onderwijsniveau in de basic skills. In de beginjaren tachtig stak een soortgelijke beweging opnieuw de kop op in de VS, toen het onderwijs in de high school onder de loupe werd genomen (Nation at Risk, 1983). Uit internationaal vergelijkend onderzoek op het gebied van wiskunde, science en talen bleek, dat de VS niet meer excelleerde – Ja-

pan daarentegen wel – en de nieuwe technologieën vroegen om beter onderwijs en in het verlengde daarvan om beter gekwalificeerd personeel.

Opmerkelijk was echter dat aan het beroepsonderwijs volstrekt geen aandacht werd besteed in Nation at Risk. Gesproken werd van de onvoltooide agenda (National Center, 1984). De roep om beter of hoger gekwalificeerd personeel wordt vertaald in programma's gericht op onderwijs in basic skills en maatregelen die de studie-inzet zowel in tijd (aantal schooldagen, huiswerk) als qua motivatie zouden versterken.

Samenvattend kan gesteld worden dat tot de basic skills traditioneel een drietal instrumentele vaardigheden worden gerekend. Zij vormen de basis 'tools' van veel leer- en communicatieprocessen in de samenleving. Het niveau waarop ze vereist worden is niet duidelijk. Door internationaal vergelijkend assessment onderzoek kunnen standaarden ontwikkeld en relatieve vergelijkingen gemaakt worden. Als de kwaliteit van het onderwijs in het geding komt, wordt de oorzaak vaak gelegd bij deze vaardigheden. De redenering is, dat wanneer deze vaardigheden ontbreken schoolverlaters niet volwaardig kunnen functioneren in de samenleving. Dat daartoe meer noodzakelijk is dan deze vaardigheden zal duidelijk zijn.

1.1.2 *Job-related basic skills*

Onder de term job-related basic skills vallen die vaardigheden die noodzakelijk zijn om adequaat te kunnen functioneren in een beroepssituatie. Dit zijn vaak, maar niet uitsluitend, technisch-instrumentele vaardigheden. In de literatuur worden drie invalshoeken gekozen van waaruit bezien wordt welke vaardigheden noodzakelijk kunnen zijn, nl.

a. het verwerven van een baan (Datta, 1982);
b. het uitvoeren van een functie (Datta, 1982; Sticht & Mickulecky, 1984);
c. het bijscholen tijdens de loopbaan (Herr, 1977).

Het verwerven van een baan is gebasseerd op een communicatieproces. Door middel van een advertentie maakt een werkgever zijn wensen bekend. De sollicitant moet deze kunnen lezen, begrijpen en interpreteren. Daarop volgt een reactie: het schrijven van een brief of het telefoneren, waarbij op duidelijke en adequate wijze de geschiktheid en motivatie voor de functie moet worden aangetoond.

Indien de situatie zover komt, dient de kandidaat doorgaans in een gesprek een en ander nader toe te lichten. Hierbij blijken houdingsaspecten zeer belangrijk.

Centraal in dit communicatieproces staan schriftelijke en mondelinge taalvaardigheden. Als een individu zijn basisvaardigheden (basic skills) niet beheerst ontstaan problemen. Het blijkt dat personen afkomstig uit etnische minderheden en uit lagere sociaal economische milieus hierbij problemen ervaren, waardoor een drempel ontstaat om een baan te verwerven (Herr, 1977).

Als een persoon de baan verwerft, ontstaat een tweede fase met voorwaarden. In welke mate en tot welk niveau reiken zijn basisvaardigheden, noodzakelijk voor de uitoefening van de begeerde functie. Sticht en Mickulecky (1984) hebben een secundaire analyse uitgevoerd op een twintigtal studies. Daaruit blijkt dat gemiddeld door technische werknemers 2 uur per dag wordt besteed aan het lezen van instructies, tabellen, kaarten en computerterminals. Deze activiteiten zijn noodzakelijk om de taken uit te kunnen voeren. Ongeveer 75% van de leesactiviteit is bedoeld om de taak uit te kunnen voeren en niet om er iets van te leren dat begrepen en onthouden moet worden (Datta, 1982; Sticht & Mickulecky,1984).

In 98% van de gevallen, zo blijkt uit hun analyse, wordt een tamelijk hoog niveau van basic skills vereist, d.w.z. een beheersing op het niveau van de tweede/derde klas voortgezet onderwijs. Sticht en Mickulecky differentiëren dit vereiste niveau helaas niet naar de soort activiteit van de arbeider. Belangrijk is hun conclusie dat de testscore die mensen behalen bij het werken met en beantwoorden van vragen over gestandaardiseerde teksten niet maatgevend is voor hun basic skills niveau op de werkplek. Vaak zijn andere cognitieve vaardigheden in het geding. Een differentiatie naar functieniveau is hier van belang.

De studie van Sticht en Mickulecky is met name gericht op uitvoerende functies. Daarbij blijkt dat een gebrek aan job-related skills tot produktieverlies of ongelukken kan leiden, omdat mensen instructies niet (goed) begrepen hebben. Voor hogere functies in de lijn of staf kan het aandeel om te begrijpen wel gro-

ter zijn. Daarover geeft de studie van Sticht en Mickulecky echter geen uitsluitsel.

Ten slotte de job-related basic skills in het kader van bijscholing tijdens de loopbaan. Een conclusie uit het voorgaande is dat als mensen al problemen hebben met basic skills, zij dit zeker zullen hebben met bijscholingsprogramma's. Herr (1977) vermeldt dat een categorie van 20 à 30% van de werknemers functioneel analfabeet is. Sticht en Mickulecky (1984) pleiten voor trainingsprogramma's met name in het basic skill domein dat sterk aan de beroepssituatie gerelateerd is. Die training zou geïntegreerd moeten worden in de training van technische beroepsvaardigheden, waardoor het nuttig effect van die trainingen oploopt tot wel drie maal dat van een normaal trainingsprogramma. In feite pleiten Sticht en Mickulecky hier voor interne bedrijfsopleidingen die functiegericht van karakter zijn met aandacht voor basic skills.

Als conclusie van deze paragraaf kunnen we formuleren dat mondelinge en schriftelijke taalvaardigheid van essentieel belang is om een baan te verwerven. Tijdens de uitoefening van de baan of functie ligt de nadruk sterk op de beoefening en het gebruik van basic skills om informatie te verwerken die direct gebruikt wordt voor de beroepsuitoefening. In het kader van de bijscholing wordt een integratie van technische en basic skills bepleit: functiegericht opleiden.

1.1.3 *Verruiming van de basic skills*

Verschillende auteurs voorzien dat ten gevolge van de invoering van nieuwe (informatie-) technologieën het niveau en de inhoud van het pakket basic skills zich zal wijzigen. Een eerste verruiming van het domein van basic skills treffen we aan in een rapport van Daniels en Karmos (1983). Zij beschrijven een onderzoek waarbij respondenten het niveau moesten aangeven van high school-studenten op de volgende basic skills domeinen: lezen, schrijven, spreken, luisteren, wiskunde, natuurkunde en redeneren. We zien in deze opsomming een uitbreiding optreden naar de meer exacte vakken in het voortgezet onderwijs, terwijl redeneren (reasoning) doelt op analytisch denken en probleemoplossen.

Mull en Prescott (1984) en Rumberger (1984) leggen vooral op deze laatste vaardigheid een zwaar accent. We komen daarop terug.

In de International Encyclopaedia of Educational Research (Husen & Postlethwaite, 1985) wordt de verwachting uitgesproken dat leerlingen en volwassenen 'media- en computer literacy' moeten leren. Onder medialiteracy wordt de vaardigheid verstaan om te leren van televisie, films en radio. Computer literacy loopt volgens hen uiteen van het kunnen interacteren met een computer d.m.v. een beeldscherm tot en met het kunnen schrijven van een computerprogramma. Mull en Prescott (1984) citeren de National Science Board Commission die stelt, dat we terug moeten keren naar de basics. De basics van de 21ste eeuw zijn echter niet meer alleen lezen, schrijven en rekenen. Deze basics omvatten nu ook communicatieve en technologische vaardigheden, denkvaardigheden waarmee we de technologische wereld om ons heen kunnen begrijpen. Rumbergers (1984) betoog gaat in dezelfde richting. Hij benadrukt dat basic skills en basisonderwijs verder reiken dan wis- en natuurkundeonderwijs; ook communicatieve, analytische en technische vaardigheden moeten getraind worden.

De invulling van het begrip basisvaardigheden, zoals die door de voorgaande auteurs is gegeven beslaat vaardigheden die een beroepsoverstijgend karakter hebben en de mogelijkheid bieden van transfer naar andere leer- en beroepssituaties. Jones (1981) legt de nadruk op 'transfer skills', omdat de enige zekerheid in de moderne samenleving de constante verandering is. Zij voert daarbij een pleidooi voor nieuwe instructiestrategieën, omdat er anders een verdere scheiding tussen de hoog- en laagpresterende leerlingen zal optreden. Als er geen nieuwe instructiestrategieën worden toegepast dan zullen de nieuwe basic skills waar de samenleving om vraagt, zoals 'planning' en 'trouble shooting' (zie b.v. Nijhof & Kommers, 1985), waarschijnlijk alleen in curricula voor hoogbegaafden worden opgenomen, vooral ook daarom omdat dit vaardigheden zijn waarmee lager begaafden vanwege het abstracte karakter altijd al moeite hebben gehad.

De conclusie die we uit deze paragraaf trekken is, dat er nog geen eenduidigheid bestaat over de verruiming van de basic skills. Wel zijn de auteurs het eens over het feit dat er een uitbreiding naast de bestaande drie 'R's' nodig is. In welke richting en binnen welk inhoudelijk domein is echter nog niet duidelijk. Wel

is duidelijk dat meer abstracte kennis en vaardigheden vereist zijn: kennis van procedures en regels en kennis van de inhoud. Dat wil zeggen contextgebonden kennis gecombineerd met leerstrategieën en andere procedurele vaardigheden.

1.1.4 *Generic skills*

De generic skills benadering verschilt in een tweetal opzichten van de hiervoor besproken benaderingen. Op de eerste plaats gaat het hierbij expliciet en uitsluitend om kennis en vaardigheden welke voor de beroepsuitoefening noodzakelijk zijn. Op de tweede plaats wordt het normatieve karakter op basis waarvan niveau en inhoud van de basic skills wordt bepaald, ingeruild voor een meer empirische benaderingswijze. We zullen aandacht besteden aan een drietal onderzoeksprojecten, waarin een en ander aan de orde wordt gesteld. Mertens'(1974) uitgangspunt is het begrip sleutelkwalificaties. Deze zijn opgebouwd uit een viertal categorieën: basiskwalificaties, horizontale of laterale kwalificaties, 'Breitenelemente' en zogenaamde vintagefactoren. In dit kader behandelen we vooral de 'Breitenelemente'. Verderop komen we op de overige categorieën terug. Mertens definieert 'Breitenelemente' als die kennis en vaardigheden, die voor een groot deel van het beroepenveld als praktische vereisten voor de invulling van de arbeidsplaats worden geformuleerd. Om deze functievereisten te achterhalen of te identificeren stelt Mertens voor functievereisten te analyseren b.v. aan de hand van beroepengidsen. Stel dat in 500 beroepen, waarvoor het onderwijs opleidt, de eindtermen meettechniek, arbeidsveiligheid en machine-onderhoud voorkomen. Mertens meent dat deze doelstellingen dan in het algemeen vormende deel van het onderwijs opgenomen zouden moeten worden. Hij kiest dus voor de grootste gemene deler van functiebestanddelen (kennis- en vaardigheidsaspecten) als criterium om te bepalen wat wel of niet tot het algemeen vormende onderwijs gerekend zou moeten worden. Hoe de 'Breitenelemente' geïdentificeerd zouden moeten worden, daarover laat hij zich niet uit.

Randhawa (1978) presenteert een methode die gebruikt is om generic skills te genereren in het kader van beroepsgericht onderwijs en training. Als uitgangspunt werd een definitie van generic skills opgesteld. Deze luidt als volgt: "Generic skills are those overt and covert behaviors which are fundamental to the performance of many tasks and subtasks carried out in a wide range of occupations and which are basic to job specific skills". Hiertoe rekent hij mathematische, communicatieve, redeneer-, interpersoonlijke en manipulatieve vaardigheden. De doelstelling was om een groot aantal beroepen te clusteren op overeenkomstige vaardigheidsprofielen. Er werd een grootschalig empirisch onderzoek uitgevoerd onder 490 werknemers en hun direct leidinggevenden. Gebruik werd gemaakt van een gestructureerde vragenlijst met 226 items verdeeld over de vijf vaardigheidscategorieën. De resultaten van de interviews werden in een beroepsvaardighedenmatrix samengebracht met voor ieder item een schaalwaarde. Het eindresultaat besloeg 164 vaardigheden. Een item kreeg een score 3 als deze noodzakelijk was voor het beroep, een score 2 als deze mogelijk noodzakelijk zou zijn, of een score 1 als deze niet vereist was voor het beroep. Op basis van deze waardering en scoring kon worden bepaald dat 38 vaardigheden tot de kern konden worden gerekend van het beroep. Deze core skills lopen uiteen van het basic skills domein, zoals eerder beschreven, naar meer beroepsgerelateerde vaardigheden, zoals telefoongebruik, het geven van informatie en instructie, of taakgerichte conversatie.

De overige 126 vaardigheden werden m.b.v. factoranalyse tot een zestal clusters samengevoegd: a. meetkunde, b. lezen en schrijven, c. algebra, d. diagnostiek, e. communicatie en f. een restcategorie.

De kernvaardigheden samen met de zes clusters kunnen gegeneraliseerd worden over de 37 onderzochte beroepen en opgenomen worden in een algemeen gedeelte van beroepsvoorbereidende scholing. Of de gevonden vaardigheden bijdragen aan een adequater en meer flexibel functioneren in de door technologische ontwikkelingen veranderende beroepssituatie kan uit de studie niet worden afgeleid, ook niet omdat onduidelijk is in welke mate moderne technologische ontwikkelingen mede bepalend zijn geweest bij de samenstelling van de itemlijst.

In Nederland wordt door het Centrum voor Innovatie van Beroep en Bedrijf (CIBB) een procedure toegepast die sterke verwantschap vertoont met het werk van Randhawa. De methode staat bekend onder de naam Mantel-

project.

Het resultaat is een aantal beroepsprofielen met een gemeenschappelijk en een gedifferentieerd deel. Het gemeenschappelijk deel kan opgevat worden als de groep generic skills voor die beroepen. Deze beroepsprofielen worden voorgelegd aan een panel van deskundigen uit de beroepssector om ze te beoordelen op volledigheid en representativiteit voor de werkelijke beroepssituatie. Een belangrijk criterium is of door de beroepsprofielen ook mogelijke toekomstige ontwikkelingen in de beroepssector worden afgedekt.

Een belangrijk verschil met de procedure van Randhawa is dat het Mantelproject per project een aantal verwante beroepen onderzoekt, welke een beperkte horizontale en verticale differentiatie kunnen vertonen. Randhawa richtte zich op een grote groep van uiteenlopende beroepen. Naast het Mantelproject bestaat een veelheid van dergelijke methoden (zie Carlisle, 1986). Doel van deze methoden is vooral de kern van de functie op te sporen, om op basis daarvan opleidingsprofielen en curricula te kunnen inrichten. Niet zozeer de vraag naar generic skills staat centraal. De gemeenschappelijke kern kan echter opgespoord worden. Belangrijk blijft daarom de vraag of d.m.v. een mantelproject-achtige werkwijze over beroepsgroepen heen generic skills kunnen worden opgespoord. In beginsel kan dit, zoals Mertens heeft getracht, doch de primaire doelstelling van dit type projecten is domein- en beroepsgroepspecifieke vaardigheidsprofielen op te sporen t.b.v. leerplanontwikkeling.

Concluderend kunnen we stellen dat de generic skills benadering een belangrijke stap voorwaarts is om inzicht in en feitenmateriaal te verkrijgen over de harde kern van kennis en vaardigheden van brede beroepsgroepen, of in meer beperkte zin voor een aantal verwante beroepen. Het zou een belangrijke stap voorwaarts zijn geweest als een sterker accent op de beroepsgebonden vaardigheden gelegd had kunnen worden.

1.1.5 *Transferable skills*

De generic skills benadering van Randhawa werd ontwikkeld in het begin van de jaren zeventig. Dan wordt door het voormalig National Center for Research in Vocational Education te Columbus, Ohio in de VS een onderzoeksproject gestart naar de relatie tussen beroepsadaptatie en transferable skills (Pratzner, 1978). 'Transferable' of 'flexible skills' zijn vaardigheden die de overgang van de ene carrière (loopbaan) naar de andere bevorderen (Wiant, 1987, p. 1). Schattingen in de VS geven aan dat men verwacht dat mensen gedurende hun leven gemiddeld 10 keer van baan zullen veranderen. 'Transition-vaardigheden zijn dan van grote betekenis. We zullen in deze paragraaf een drietal benaderingen bespreken die alle min of meer pogen om de relatie tussen beroepsadaptatie en flexibiliteit of transferability aan de orde te stellen, dan wel te onderzoeken. Op de eerste plaats Daniels en Karmos (1983) die een drietal alternatieven beschrijven om studenten te scholen in het adapteren van nieuwe technologische ontwikkelingen in de samenleving. In de tweede plaats het onderzoeksproject Transferable skills van het voormalig National Center for Research in Vocational Education te Columbus, Ohio (Pratzner, 1978; Wiant 1987). In de derde plaats het onderzoek naar sleutelkwalificaties door Mertens (1974,1975).

1.1.5.1 *Daniels & Karmos' adaptatieconcept*

Het onderzoek van Daniels en Karmos is gericht op de vraag hoe studenten voorbereid kunnen worden voor een samenleving die in hoge mate door technologie wordt beïnvloed. Zij presenteren drie alternatieven die als interventie- of instructiestrategieën kunnen worden opgevat. De drie alternatieven bouwen op elkaar voort en hebben daardoor een cumulatief karakter:

a. het belang van basic skills. Het gaat hierbij om de bekende basisvaardigheden als lezen, schrijven, spreken, wiskunde, luisteren, natuurkunde, redeneren.
b. Het tweede alternatief legt naast de basic skills ook een accent op transferable skills. De laatste worden als volgt gedefinieerd:

1. de vaardigheid is fundamenteel voor een bepaald beroep of trainingsprogramma; (entreevoorwaarde)
2. de vaardigheid is noodzakelijk voor het succesvol kunnen uitvoeren resp. volgen van een beroep of trainingsprogramma; (handhavingsconditie)
3. de vaardigheid is overdraagbaar naar andere beroepssituaties en -categorieën (transitievoorwaarde). Op grond van nadere analyse formuleren Daniels en Karmos

1. Herkennen van en reageren op veranderingen; een persoonlijke stijl van reageren ontwikkelen.
2. Taxeren en gebruiken van externe symptomen; gebruik maken van externe ondersteunende systemen, zoals gezin, vrienden, trainingsprogramma's.
3. Taxeren en gebruiken van interne symptomen; gebruik maken van het persoonlijke bewustzijn als prikkel tot verandering en persoonlijke kracht, als een bron van energie voor verandering.
4. Verminderen en kunnen hanteren van emotionele spanning.
5. Plannen en uitvoeren van veranderingen; gebruik maken van discrepanties tussen bestaande en gewenste omstandigheden om op basis daarvan een actieplan op te stellen, om een nieuwe visie te ontwikkelen en uit te voeren.

Figuur 1 *Change skills*

een vijftal categorieën transferable skills: wiskundige, communicatieve, interpersoonlijke, redeneer- en manipulatieve vaardigheden.

c. Het derde alternatief bouwt voort op de voorgaande twee en omvat derhalve basic skills en transferable skills. Daniels en Karmos voegen daar een pakket 'change skills' aan toe. Deze vaardigheden om te kunnen veranderen vormen de basis voor een adequaat zelf-management. De auteurs menen dat het derde alternatief de beste mogelijkheid biedt om in te spelen op de veranderende eisen van de moderne technologische samenleving.

Om de genoemde change skills te concretiseren geven we van hen in Figuur 1 een overzicht.

In een latere publikatie definiëren Daniels, Karmos en Presley (1985) de vaardigheden onder andere labels, zoals generalizable skills, problem solving skills en transition skills and knowledge. Waarmee aangegeven wordt hoe moeizaam het definitieprobleem rondom het begrip basisvaardigheden wordt opgelost.

1.1.5.2 *Het Transferable Skills Project*

Loopbanen van mensen voltrekken zich niet meer zoals decennia geleden. Meerdere carrières in een leven zal steeds meer usance worden. Dit gegeven vormde de achtergrond van het project Transferable Skills. De vraag die op kwam was welke vaardigheden dragen ertoe bij dat de overgang van de ene carrière naar de ander wordt bevorderd en de vraag waar en hoe deze te leren zijn. Wat flexibele of transferable skills waren was aanvankelijk niet zo duidelijk "(...) imprecisely defined skills such as communication, interpersonal skills, problem solving skills and so forth" (Wiant, 1987, p.2); alhoewel eenieder die bij het project betrokken raakte, overtuigd was van het belang van deze vaardigheden.

Volgens Pratzner (1978) kunnen transferable skills in twee betekenissen worden opgevat. Ten eerste als vaardigheden waardoor individuen aangeleerde of bestaande kennis en vaardigheden kunnen overdragen en toepassen in nieuwe of veranderende situaties. Ten tweede als vaardigheden met transferwaarde. Dit zijn vaardigheden waarvan gebleken is dat ze overdraagbaar zijn naar vele (nieuwe) vaardigheden. Het is een belangrijk analytisch onderscheid en roept twee achterliggende vragen op: Wat zijn transfervaardigheden die het transferproces bevorderen, en wat zijn transferable vaardigheden?

We gaan in deze paragraaf op de tweede vraagstelling in, ook omdat deze de kern vormt van het Transferable skills project. Om de vraag naar transferable skills te kunnen beantwoorden werd een inductieve procedure gevolgd. Veertien onderwijsprogramma's van scholen en bedrijven werden geselecteerd. De programma's werden onderzocht op impliciete en expliciete aspecten van vaardigheden die transfer bevorderen. Daarnaast werden bijeenkomsten belegd met docenten, vakbondsvertegenwoordigers, personeels- en opleidingsfunctionarissen uit verschillende bedrijfstakken.

De deelnemers werd verzocht om hun ervaringen en meningen over tranfervaardigheden, personeelsselectie, training, feitelijke transfer van personen en functiegroei te vertellen. De onderzoekers benadrukken dat de resultaten die deze analyses opgeleverd hebben, geen brede empirische basis hebben en daarom met de nodige voorzichtigheid moeten worden gehanteerd. De resultaten van de analyses en groepsinterviews werden verwerkt in een vijftal categorieën, en wel:

a. mathematische vaardigheden,
b. communicatieve vaardigheden,
c. interpersoonlijke vaardigheden,

d. redeneervaardigheden,
e. psycho-motorische vaardigheden.

De beschreven vaardigheden zijn zeer algemeen van karakter, fundamenteel voor vele beroepsgroepen of opleidingsprogramma's, noodzakelijk voor succes en overdraagbaar naar andere beroepssituaties. Door Pratzner worden deze vaardigheden als noodzakelijk genoemd in verband met de door de technologieën veranderde beroepspraktijk. Deze inductieve inventariserende benadering van Pratzner geeft geen duidelijkheid over de richting van transfer; er wordt bijvoorbeeld geen uitsluitsel gegeven op de vraag of het om horizontale of verticale transfer gaat, of het om 'near' dan wel 'far' transfer handelt.

Pratzner plaatst wel kanttekeningen bij dit soort lijsten met vaardigheden. Er is, zo stelt hij, erg weinig bekend over functieveranderingen en beroepsmobiliteit. Er is geen algemeen aanvaarde lijst met specifieke vaardigheden en persoonskenmerken die ingezet kunnen worden in een breed scala aan beroepssituaties. Transfervaardigheden als zodanig garanderen nog geen groter adaptatievermogen; zij bevorderen deze hoogstens. Overigens zijn de laatste jaren wel studies ondernomen naar loopbanen (Balmer, Gonon & Straumann, 1986). Vooral de biografische methode is daarin zeer in zwang. Wij komen daarop bij de beschrijving van Mertens terug.

1.1.5.3 *Sleutelkwalificaties: Het concept van Mertens*

Het begrip sleutelkwalificaties dat door Mertens (1974) is ontwikkeld, komt voort uit een onderzoeksprogramma van het Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB) te Nürnberg (BRD). Het onderzoek startte in de jaren zestig en probeerde een oplossing te vinden voor de starheid van de onderwijseconomische planningsmodellen. Met behulp van een flexibiliteitsonderzoek is getracht hierop een antwoord te vinden. Getracht wordt vast te stellen in welke mate personen met dezelfde opleiding of hetzelfde beroep in uiteenlopende functies te vinden zijn en in hoeverre bepaalde soorten functies door personen met uiteenlopende opleidingen of beroepen worden bezet. Zo hoopt men flexibiliteitsmogelijkheden aan de aanbodzijde en de vraagzijde op het spoor te komen (Dedering & Schimming, 1984; Van Hoof & Dronkers, 1980). Het flexibiliteitsonderzoek heeft een sterk empirisch karakter. Het begrip sleutelkwalificaties en de nadere uitwerking daarvan is echter normatief van aard en vormt een aanvulling in onderwijskundige richting op dit arbeidsmarktonderzoek.

In het flexibiliteitsonderzoek worden substitutie- en mobiliteitskenmerken van de arbeidsmarkt onderzocht. In deze benadering worden nog geen ontwikkelingen in de beroepssituatie en de werkgelegenheidsstructuur verdisconteerd. Sleutelkwalificaties dienen zodanig van karakter te zijn dat ze maximale substitutie en mobiliteit garanderen en in potentie aan toekomstige, gedeeltelijk nog onbekende kwalificatievereisten voldoen. Mertens geeft in feite een aanzet tot een taxonomie van sleutelkwalificaties met een heuristische waarde. Sleutelkwalificaties dienen categorieën te bevatten m.b.t. de bevordering van een veelheid aan vaardigheden en persoonskenmerken. Hij noemt als voorbeelden: het levenslang kunnen leren, wisseling van sociale rollen, afstand scheppen door te theoretiseren, creativiteit en relativeringsvermogen bevorderen, verbinden van theorie en praktijk, ontwikkelen van technisch begrip, analyse van interesses, samenwerkingsbereidheid, uithoudingsvermogen, concentratie, etc.

De taxonomische indeling van Mertens bestaat uit 4 categorieën:
a. basiskwalificaties, kwalificaties van een hogere orde, gericht op verticale transfer;
b. horizontale kwalificaties, kwalificaties gericht op horizontale of laterale transfer (soortgelijke situaties, functies, beroepen);
c. 'Breitenelemente' kunnen worden opgevat als generieke vaardigheden (zie ook paragraaf 3.2.4);
d. 'vintage factoren' omvatten die kennis en vaardigheden die in de huidige opleidingen wel zijn opgenomen, maar die oudere generaties niet meer hebben gehad (vreemde talen, computer onderwijs, e.d.). Zij kunnen als inbreng dienen voor volwasseneneducatie of voor bijscholingscursussen.

In het kader van deze paragraaf zijn de basiskwalificaties en de horizontale kwalificaties van betekenis, daar zij immers verwijzen naar transfer.

Basiskwaliteiten en horizontale kwalificaties kunnen worden onderscheiden in een aantal deelvaardigheden (zie Figuur 2).

Basiskwalificaties	Horizontale kwalificaties	
algemeen	*algemeen*	*concretisering*
1. logisch denken	1. geïnformeerd zijn over informatie	a. het wezen van informatie
2. analytisch denken		b. inwinnen van informatie
3. kritisch denken		c. begrijpen van informatie
4. structurerend denken		d. verwerken van informatie
5. dispositie denken		
6. coöperatief denken		
7. conceptueel denken		
8. beslissingsgericht denken		
9. creatief denken		
10. contextueel denken		

Figuur 2 *Sleutelkwalificaties volgens Mertens* (1974)

Door Balmer, Gonon en Straumann (1986) is een biografisch onderzoek uitgevoerd naar innovatie en kwalificatie van elektromonteurs en mecaniciens als gevolg van nieuwe technologieën. Hun studie was erop gericht onder andere na te gaan in welke mate kwalificaties doorwerken van school naar beroep en binnen verschillende functies (carrière). Het kwalificatiemodel dat zij schetsen is een leerproces op de arbeidsplaats zelf dat zij een immanent karakter toekennen. De speelruimte die iemand krijgt om zichzelf te ontwikkelen wordt bepaald door drie dimensies: de mate van generaliseerbaarheid van kwalificaties, de mate van zelfstandigheid en de individuele leerstrategie van de werknemer.

Waar het de generaliseerbaarheidsthese of -dimensie betreft sluiten de auteurs aan bij Mertens' conceptie om deze empirisch te onderbouwen. Anderzijds menen zij wel dat Mertens' polyvalente kwalificaties zo algemeen en situatieneutraal zijn dat zij moeilijk op concrete beroepssituaties betrokken kunnen worden.

Uit hun analyse van biografieën blijkt dat ook heel duidelijk t.a.v. zowel de 'Breitenelemente' als t.a.v. basiskwalificaties. Balmer et al. (1986, p.166) gaan zelfs zo ver dat zij de basiskwalificaties als zelfstandig definieerbare kwalificaties afschrijven. Zij menen dat het vooral *socialisatieprocessen* zijn die transfer bewerkstelligen op de werkvloer. Logisch denken b.v. is een eis die altijd wel geformuleerd kan worden bij welk beroep dan ook. De toepassing van logisch denken zal bij een informaticus verschillen van die van een verpleegkundige evenals de wijze waarop deze wordt geleerd. Generalisaties in de vorm van specifieke socialisatie-effecten zijn door Balmer et al. duidelijk geconstateerd. Ook Wiant (1987) wijst in zijn bijdrage meer op de wijze van verwerven van flexibele vaardigheden, dan dat hij zich vermoeit met de vraag wat ze zijn.

Wie de vaardigheden van Mertens bekijkt en nader analyseert zal moeten vaststellen dat deze inderdaad zeer algemeen zijn en van een hoog abstractieniveau. Deze abstractiegraad vloeit voort uit de eis van algemene wendbaarheid, inzetbaarheid en toepasbaarheid. Juist t.a.v. deze abstractiegraad formuleren verschillende auteurs zware kritiek op het sleutelkwalificatieconcept van Mertens. De kwalificaties zijn veel te abstract, staan te ver af van de concretisering in de onderwijspraktijk en zeker van die in het beroepsonderwijs. De kwalificaties zijn zowel te vaag als te ruim, ze zullen nader geoperationaliseerd moeten worden. Dan kan ook de feitelijke transferpotentie onderzocht worden (Dedering & Schimming, 1984; Balmer et al.,1986).

Op basis van het voorgaande kunnen we enkele conclusies trekken. De empirische basis waarop de uitspraken en voorstellen voor transfervaardigheden berusten is nogal zwak. De gegevens moeten daarom met het nodige voorbehoud worden gehanteerd. De genoemde vaardigheden zijn van een buitengewoon hoog abstractieniveau, zodat ze nauwelijks differentiëren tussen beroepen en soorten onderwijs.

Veel vaardigheden die genoemd worden bevinden zich in het klassieke cognitieve instrumentele domein. Dat is op zich niet verwonderlijk. Daarentegen valt wel op dat

slechts marginaal aandacht bestaat voor niet-instrumentele kwalificaties, ook wel houdingsaspecten genoemd. Daarnaast kan gevraagd worden in hoeverre cognitieve bekwaamheid de anderssoortige vaardigheden (interactief, psychomotorisch) beïnvloedt. Ten slotte valt wel op dat er een grote mate van overeenstemming lijkt te bestaan in de drie geschetste benaderingen. Er worden op drie momenten in de carrière van het beroepsverloop getracht keuzen te maken: bij de entree (basisvaardigheden), bij de uitvoering en handhaving van de functie (generic and job-related skills) en bij de transitie van de ene functie of carrière naar de andere. Het accentueren van change skills/flexible skills om mee te kunnen groeien met de arbeidsorganisatie, of om het beroepsleven zelf een andere wending te kunnen geven (self management) is een opvallend element in het werk van Daniels, Karmos en Presley (1985). Vanuit een tamelijk sceptische instelling kunnen we ons afvragen of bijna alle genoemde vaardigheden transferpotentie bezitten? Wiant (1987) geeft zelf een positief antwoord op die vraag en hij stelt ook nuchter vast dat de identificatie van zo algemene vaardigheden geen 'nieuwe ontdekking' is.

2 *Naar een voorlopige synthese*

We hebben verschillende auteurs de revue laten passeren. Het was de bedoeling om tot dieper inzicht te komen, de essentie van het begrip basisvaardigheden in de greep te krijgen en zo mogelijk tot een nadere concrete invulling te komen. Een vijftal optieken is beschreven. Het is nu de bedoeling om deze zo mogelijk te verbinden of te synthetiseren.

We hebben in het voorgaande gezien dat het begrip basisvaardigheden wordt geënt op de beroepsrol, soms heel specifiek in relatie tot (nieuwe) technologische ontwikkelingen. Daarbij vallen drie elementen te onderscheiden:

a. de *entree* op de arbeidsmarkt en de noodzakelijke beroepsvoorbereiding daarvoor. Men dient te beschikken over bepaalde (beroeps- of functiegebonden) basisvaardigheden: (basic skills/job-related skills);
b. de *uitvoering* van het beroep (het handhaven en verder doorgroeien in de functie; (generic skills/job-related skills)
c. de *overgang* van de ene functie naar de andere in een soortgelijk beroep, of de overstap (transitie) van carrière naar carrière (flexibel skills/transition skills/change skills).

Men zou de vaardigheden ook anders kunnen ordenen en wel als volgt:

1. *Voorwaardelijke vaardigheden*
   Deze behoren tot het domein van de basisvorming voor iedereen. De WRR (1986) definieert basisvorming als "de gemeenschappelijke en algemene vorming op intellectueel, cultureel en sociaal gebied, die als grondslag dient voor een verdere ontwikkeling van de persoonlijkheid, alsmede voor een zinvol functioneren als lid van de samenleving en voor een verantwoorde keuze en een verdere scholing van een beroep." Voor zover voorbereiding voor een beroep plaatsvindt zullen basic skills noodzakelijk zijn aangevuld met oriënterende job-related skills.
2. *Gemeenschappelijke beroepsvaardigheden*
   De vaardigheden die hier beoogd worden zijn beroepsgebonden vaardigheden en vallen uiteen in een tweetal categorieën: a) vaardigheden die gemeenschappelijk zijn voor *alle* beroepen; en b) vaardigheden die gemeenschappelijk zijn voor *bepaalde* beroepsgroepen.
3. *Transitie-vaardigheden*
   Deze vaardigheden zijn van een hogere orde en stellen een persoon in staat om informatie uit de beroeps- en arbeidswereld zodanig te verwerken dat een adequaat besluitvormingsproces kan plaatsvinden om een carrière te plannen resp. uit te voeren (self management; leren leren).

De *voorwaardelijke vaardigheden* worden gevormd door onder andere de basic skills, die overigens meer omvatten dan lezen, schrijven en rekenen. De standaard voor een moderne samenleving bestaat ons inziens minimaal uit vreemde talen onderwijs, moedertaal, wiskunde, science, techniek en technologie, social studies, expressievakken, beroepenoriëntatie, economie en politiek maatschappelijke vorming. Iedereen moet een minimale beheersing vertonen op deze domeinen om in de samenleving te kunnen functioneren. Ze vormen de voorwaarde om de gemeenschappelijke en transitievaardigheden te leren.

In zekere zin geldt dit ook voor de job-

related- en de verruimde basic skills. Zij nemen in ons betoog een tussenpositie in. Lezen, schrijven en rekenen vormen voorwaarden voor communicatie, voor media en computer literacy, voor wiskunde en natuurkunde. Toch is de veronderstelling dat al deze vaardigheden een dusdanig algemeen karakter hebben, dat zij voorwaardelijk zijn voor het functioneren van elk individu in onze huidige samenleving. Beroepsgebonden of specifiek zijn ze daardoor niet.

De *generic skills* vormen een pendant van de gemeenschappelijke beroepsvaardigheden. Door middel van statistisch-analytische methoden wordt van een groot aantal beroepen bepaald wat het gemeenschappelijk deel aan kennis en vaardigheden is. Door het feit dat ze gemeenschappelijk zijn is een uitwisseling mogelijk en bezitten deze vaardigheden transferpotentie. Toch lijkt het ons niet gewenst om te spreken van basisvaardigheden in de zin van basisvorming. Het gaat hier eerder om 'common', d.i. gemeenschappelijke elementen, dan om core skills in de betekenis van fundament voor andere vaardigheden. Beroepsspecifieke training maakt wel gebruik van basic skills, maar niet altijd van generic skills.

De *transitie-vaardigheden* zouden aan dit laatste probleem tegemoet moeten komen. De transferable of flexible skills vormen deze laatste categorie. Het is een belangrijke maar nog moeilijk grijpbare categorie. Balmer et al. hebben gewezen op het generalisatie-effect dat optreedt door socialisatie, maar tevens op het zelfstandigheids- en autonomie-aspect in een beroep en de leerstrategie waarover iemand beschikt. We kunnen deze vaardigheden dus zien als resultaat van een bepaald leer- of vormingsproces en als voorwaarde voor verdere ontwikkeling en vorming. Juist dit voorwaardelijk karakter maakt ze erg belangrijk in het kader van her- en bijscholing en voor adaptatie in veranderende beroepssituaties.

De analyse tot nu toe leert ons dat we de groepen basisvaardigheden hebben uiteengelegd in drie categorieën. We hebben een verbinding gemaakt met faseringen in de arbeidsloopbaan, waardoor een matrix met 9 cellen kan worden gevormd (zie Figuur 3).

Vaardigheden / Fasen	Basic	Generic	Transferable
entree	x	x	
uitvoering volharding		x	x
meervoudige carrières		x	x

Figuur 3 *Vaardigheden gerelateerd aan loopbaanfasen*

Elke groep van vaardigheden zou vervolgens geinventariseerd en geanalyseerd kunnen worden in onderwijskundig taxonomische termen. Daaruit komt dan een specifiek resultaat dat voor een beroepsopleiding of functiegerichte training betekenisvol is. Voor een voorbeeld kan verwezen worden naar Olbrich en Pfeiffer, 1980; voor een toepassing zie Nijhof en Mulder, 1986. Het belangrijkste van het schema in Figuur 3 is dat zichtbaar wordt gemaakt dat de definitie van basisvaardigheden aangescherpt moet worden en wellicht dat de term zelf moet vervallen voor het beroepsonderwijs en de volwasseneducatie. We zouden de term basisvaardigheden kunnen reserveren voor basic skills (basisvorming), terwijl voor de beroepsvoorbereiding en -training en socialisatie gedacht zal moeten worden aan generieke vaardigheden resp. flexibiliteits- of transitievaardigheden. Daarmee zijn dan resp. vaardigheden bedoeld die de laterale of horizontale transfer in beroepssituaties dienen, en verticale transfer in loopbanen waardoor de mobiliteit van werknemers kan worden bevorderd.

In de geschetste vijf benaderingen van het begrip basisvaardigheden keren standaard drie groepen inhoudelijke aspecten terug: kennisaspecten (concrete aan kennisdomeinen gebonden inhoudelijke kennis) en (psychomotorische) vaardigheden, procedurele kennisaspecten (weten hoe), en persoonskenmerken (leerstijlen, autonomie, creativiteit). Deze drie groepen treffen we aan in de drie onderscheiden typen basisvaardigheden. *Basisvaardigheden kunnen we terugkerend naar onze werkdefinitie daarom herformuleren als die instrumentele en niet-instrumentele kennis en*

vaardigheidsaspecten die verticale en horizontale transferwaarde hebben voor gediplomeerde schoolverlaters en beroepsbeoefenaren in die zin dat zij toegang verschaffen tot de beroepswereld, tot verdere studie of scholing en meervoudige carrières mogelijk maken.

3 *Discussie*

De analyse van het begrip basisvaardigheden laat zien dat ondanks verschillen in aanpak de uitkomsten heel algemeen zijn: nogal abstracte categorieën vaardigheden. Dat roept alle bezwaren op die ook bij Mertens geformuleerd zijn. Bovendien leidt die kritiek tot de vraag of er wel verschil is tussen beroepsonderwijs en algemeen vormend onderwijs. Of genuanceerder: hoe algemeen wordt beroepsonderwijs als zij meer op die algemene basisvaardigheden wordt toegespitst? Het zal duidelijk zijn dat dat onderscheid volledig verdwijnt indien uitsluitend op deze algemene vaardigheden wordt geschoold zonder enige vorm van beroepsgebonden context. De vraag is dan wel hoe zinvol en doelmatig dat is. Voor ieder zal duidelijk zijn dat metselen een *specifieke* beroepsvaardigheid is. Duidelijk is ook dat logisch denken een *algemene* vaardigheid is. Een beroepsgerichte voorbereiding maakt het noodzakelijk algemene en beroepsspecifieke vaardigheden te combineren, zoals dat in het hoger onderwijs ook het geval is. Zeer specifieke vaardigheden worden in de praktijk geleerd.

De vraag wordt daarom veeleer hoe beroepsspecifiek een bepaalde algemene vaardigheid moet worden aangekleed om functioneel te zijn voor de beroepsuitoefening. De mate en de verhouding waarin substantiële kennis gecombineerd moet worden met procedurele kennis bepaalt hoe specifiek een vaardigheid voor een bepaald beroep is.

Dat leerprocessen contextgebonden zijn is de laatste jaren met grote evidentie aangetoond. Taak- en beroepenanalyses zijn daarom noodzakelijk om uit te kunnen maken hoe algemeen of specifiek bepaalde vaardigheden zijn en hoe en waar ze geleerd of getraind moeten worden. De training van operators, loodsen en piloten is daarop b.v. gebaseerd.

Dat roept dus de vraag op of het nader operationaliseren van die algemene vaardigheden inderdaad leidt tot verschillende inhouden voor verschillende doelgroepen. Die vraag is ook betekenisvol in het kader van schaalvergrotingsprojecten zoals de SVM-operatie.

Door samenvoeging van scholen ontstaan niet alleen grote schoolgemeenschappen, maar ook nieuwe mogelijkheden om modulen uit te wisselen tussen verschillende opleidingstrajecten: Engels is immers Engels, en natuurkunde is natuurkunde? Maar is handels Engels (MEAO) hetzelfde als technisch Engels (MTO)? Is toegepaste natuurkunde hetzelfde als fundamentele natuurkunde. Verschillende doelgroepen zullen verschillende curricula behoeven, afhankelijk van de specifieke behoeften op de arbeidsplaats.

Een taakverdeling tussen of integratie van algemeen vormend en beroepsgericht onderwijs op grond van deze analyse van basisvaardigheden alleen is niet toereikend. Daarvoor is een heroriëntatie op de opleidingsstructuur noodzakelijk in het licht van de nieuwe eisen die de samenleving stelt aan haar onderwijssysteem. Dat kan betekenen dat bedrijfsopleidingen de specifieke functie van functiegerichte training opnieuw versterken. Bij grote bedrijven is dit reeds het geval. Daardoor komt in het onderwijs ruimte voor het versterken van de algemene component. Het vinden van de optimale koppeling tussen onderwijssysteem en bedrijfsopleidingen zal een belangrijke opgave zijn voor de jaren negentig.

Literatuur

Balmer, K., Ph. Gonon & M. Straumann, *Innovation und Qualifikation*. Eine berufbiografische Untersuchung von FEAM und Mechanikern. Zürich: Institut für Bildungsforschung und Berufspädagogik des kantonalen Amtes für Berufsbildung, 1986.

Carlisle, K.E., *Analyzing jobs and tasks*. Englewood Cliffs, N.J: Educational Technology Publications, 1986.

Datta, L., Employment related basic skills. In: H.F. Silberman, H.F. (Ed.), *Education and work*. Chicago: The University of Chicago Press, 1982.

Daniels, M. H. & J.S. Karmos, *Preparing Students to adapt with new technology* Carbondale, Ill.: College of Education, Southern Illinois University, 1983.

Daniels, M. H., Karmos, J. S. & C. A. Presley, *Toward Excellence in Secondary Vocational Education: Developing Pretechnical Curricula*. Columbus, Oh: The National Center for Research in Vocational Education (Information series no.295).

Dedering, H. & P. Schimming, *Qualifikationsforschung und Arbeitsorientierte Bildung*. Wiesbaden: Westalt, 1984.

Herr, E. L., *Links among Training, Employability and Employment*. Columbus, Oh.: The National Center for Research in Vocational Education, 1977.

Hoof, J. J. van & J. Dronkers, *Onderwijs en Arbeidsmarkt: een verkenning van de relaties tussen onderwijs, arbeidsmarkt en arbeidssysteem*. Deventer: Van Loghum Slaterus, 1980.

Husen, T. & T. N. Postlethwaite (Eds.), *The international Encyclopaedia of Education: research and studies*. Oxford: Pergamon Press, 1985.

Jones, K. S. (Ed.) *Information retrieval experiment*. London: Butterworths, 1981.

Mertens, D., Schlüsselqualifikationen. Thesen zur Schulung für eine moderne Gesellschaft. *Mitt. a.d. Arbeidsmarkt- und Berufsforschung*. 1974, no 7, 36-43.

Mertens, D., Schlüsselqualifikationen und Berufsbildung; Versuch einer Erwiderung. *Berufsbildung in Wissenschaft und Praxis*, 1975, no 5, 24-25.

Mull, P. M. & C. A. Prescott, *High technology careers* Waco: Dallas, Tx: The Center for Occupational Research and Development., 1984.

Nation at Risk The Commission on Excellence in Education. Washington, DC., 1983.

National Center for Research in Vocational Education, *The unfinished agenda: The role of vocational education in the high school*. Columbus: Ohio, 1984.

Nijhof, W. J. & P. Kommers, An analysis of cooperation in relation to cognitive controversy. In R. Slavin et al. (Eds.), *Learning to cooperate, cooperating to Learn*. (pp 125-147). London: Plenum Press, 1985.

Nijhof, W. J. & M. Mulder (Red.) *Basisvaardigheden in het Beroepsonderwijs*. Den Haag: SVO/TO, 1986.

Nijhof, W. J. & J. L. M. Remmers, *Basisvaardigheden nader bekeken*. Enschede: Faculteit der Toegepaste Onderwijskunde, 1989.

Olbrich, G. & V. Pfeiffer, Darstellung und Anwendung eines Hierarchisierungssystems für Lernziele in der Beruflichen Bildung. *Berichte zur Beruflichen Bildung*. Heft 25. Berlin: Bundesinstitut für Berufsbildung, 1980.

Pratzner, F. C., *Occupational adaptability and transferable skills*. Columbus, Oh.: the National Center for Research in Vocational Education. (Bulletin no.29), 1978.

Randhawa, B. S., Unstering of skills and occupations: a generic skills approach. *Journal of Vocational Behavior*, 1978. no 12, 80-92.

Rumberger, R., *Demystifying High Technology*. Columbus, Oh: National Center for Research in Vocational Education, 1984.

Sticht, T. G. & L. Mickulecky, *Job related basic skills: cases and conclusions*. Columbus, Oh: The National Center for Research in Vocational Education, 1984.

Stout, V. J. & H. W. Taylor, *Skills need for Information Technology users*. Athene, Ga: The University of Georgia, 1984.

Wegmann, R., R. Chapman & M. Johnson, *Work in the New Economy*. Indianapolis: JIST, 1989.

Wiant, A. A.,*Flexible skills and training strategies*. Columbus, Oh: The National Center for Research in Vocational Education, 1987.

WRR, *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij, 1986.

Curricula vitae

W. J. Nijhof (1941), hoogleraar Onderwijskunde, Vakgroep Curriculumtechnologie, Universiteit Twente. Onderzoeksterrein: functie- en beroepsgerichte educatie, bedrijfsopleidingen en curriculumtechnologie.

Adres: Postbus 217, 7500 AE Enschede.

J. L. M. Remmers (1958), verbonden aan het IVA (Instituut voor Sociaal Wetenschappelijk Onderzoek) als wetenschappelijk onderzoeker bij de sectie Onderwijs en Cultuur. Onderzoeksterrein: aansluiting onderwijsarbeid, de implementatieaspecten van computergebruik in de schoolorganisatie, de invoering van formatiebudgettering in het primair en voortgezet onderwijs.

Adres: IVA, Postbus 90153, 5000 LE Tilburg.

Manuscript aanvaard 4-5-'90

Summary

Nijhof, W.J. & J.L.M. Remmers. 'Transferable skills: toward a definition and operationalisation.' *Pedagogische Studiën*, 1990, *67*, 389-402.

New technologies have impact on jobs. But the direction of this impact is unclear. Nevertheless it is obvious that policy institutions like departments of education would like to know what kind of qualifications will be necessary in the near future. Especially in The Netherlands, with an education system split up into a general and vocational track, it might be interesting to know how to divide or to integrate the training system from/into the general education system part.

In this article an analysis has been made of international literature in order to find and formulate common components of a definition of transferable skills.

Ouderlijke opvoedingsgedragingen, empathie en prosociale ontwikkeling van het kind

J.M.A.M. JANSSENS, J.R.M. GERRIS, A.W.H. JANSSEN
Vakgroep Algemene Pedagogiek Katholieke Universiteit Nijmegen

Samenvatting

In dit artikel worden relaties onderzocht tussen ouderlijk opvoedingsgedrag, empathie en de prosociale ontwikkeling van het kind. Op de eerste plaats blijkt dat de prosociale ontwikkeling van het kind positief samenhangt met ouderlijke ondersteuning, met inductief disciplineringsgedrag en met gedrag waarbij de ouder een beroep doet op de verantwoordelijkheid van het kind. Machtsuitoefening hangt daarentegen negatief met de prosociale ontwikkeling van het kind samen. Op de tweede plaats blijkt dat de genoemde opvoedingsgedragingen niet onafhankelijk zijn van elkaar. Inductie en het beroep doen op de verantwoordelijkheid van het kind hangen positief samen en beide gedragingen zijn negatief gerelateerd aan machtsuitoefening. Een derde conclusie is dat empathie een positieve invloed heeft op de prosociale ontwikkeling van het kind. Bovendien is empathie afhankelijk van een opvoedingspatroon dat gekenmerkt wordt door weinig machtsuitoefenend en veel inductief gedrag en gedrag waarbij de ouder het kind aanspreekt op zijn verantwoordelijkheid. Empathie kan derhalve gezien worden als een interveniërende variabele tussen opvoedingsgedrag van ouders en de prosociale ontwikkeling van het kind.

De resultaten zijn gebaseerd op een studie onder 120 Nederlandse gezinnen met een kind tussen 9 en 12 jaar oud.

1 *Inleiding*

De laatse decennia is er veel onderzoek gedaan naar de prosociale ontwikkeling van kinderen en naar de wijze waarop ouders die ontwikkeling kunnen beïnvloeden. Wat prosociale ontwikkeling betreft wordt in dit artikel een onderscheid gemaakt tussen prosociaal moreel redeneren en prosociaal gedrag. Prosociaal moreel redeneren heeft betrekking op redeneren in situaties waarin de eigen behoeften in conflict zijn met die van anderen; de persoon geeft aan waarom aan eigen of aan andermans behoeften voorrang moet worden gegeven. Prosociaal gedrag betreft gedrag waarbij men iets doet voor een ander en daaraan eigen belangen ondergeschikt maakt.

1.1 *Disciplinering en prosociale ontwikkeling*

Hoe kunnen ouders prosociaal gedrag en prosociaal redeneren bij hun kind bevorderen? Door Hoffman (1963, 1970, 1975, 1982) is in dat verband vooral accent gelegd op de kwaliteit van disciplineren door ouders. Disciplineren is een vorm van controle-gedrag. Door het toepassen van een bepaalde disciplineringstechniek proberen ouders het gedrag van hun kind te veranderen in een door hen gewenste richting (zie voor een meer uitgebreid overzicht van het controle- en disciplineringsconcept Gerris & Janssens, 1987). Rollins en Thomas (1979) en Hoffman (1970) onderscheiden drie vormen van controle of disciplineren: machtsuitoefening, inductie en liefdesonthouding. In het geval van machtsuitoefening oefent de ouder druk uit op het kind om zich in overeenstemming met de wensen van de ouder te gedragen en de eigen behoeften ondergeschikt te maken aan die van de ouder. Voorbeelden daarvan zijn: fysieke en niet-fysieke straf of daarmee dreigen en verbaal geweld. In het geval van inductie probeert de ouder het kind te overreden zich in overeenstemming met de wensen van de ouder te gedragen. De ouder vermijdt daarbij een direct belangenconflict, maar probeert de redelijkheid van zijn verzoek te doen inzien. Voorbeelden zijn: uitleg geven, redenen geven voor gewenst gedrag, wijzen op de gevolgen van het gedrag voor het kind zelf of voor anderen. In het geval van liefdesonthouding wijst de ouder het gedrag van het kind af en dreigt pas

en dreigt pas dan de relatie te herstellen als het kind tot inkeer komt. Voorbeelden zijn: het kind negeren, isoleren, teleurstelling uitspreken en koel reageren op het kind. Volgens Hoffman (1975) zou inductie positief en zou machtsuitoefening negatief samenhangen met prosociaal gedrag, terwijl voor liefdesonthouding de samenhang met prosociaal gedrag zou ontbreken. Op de reden daarvan zijn we elders uitvoerig ingegaan (Janssens & Gerris, 1984; Gerris & Janssens, 1984). Kort samengevat komt de theorie van Hoffman er op neer dat de positieve invloed van inductie op prosociale ontwikkeling gemedieerd wordt door empathie. Daaronder wordt het zich kunnen inleven in en meeleven met anderen verstaan. Volgens Hoffman en Saltzstein (1967) motiveert inductie het kind zich te oriënteren op de schade of het leed die het veroorzaakt heeft bij een ander. Het besef dat men de oorzaak is van dat leed of die schade kan empathische gevoelens doen onstaan. Men leeft mee met het slachtoffer. Dat in- en meeleven kan op zijn beurt weer aanleiding geven om de schade te herstellen of het leed van het slachtoffer te verzachten. Machtsuitoefening en liefdesonthouding kunnen daarentegen ressentiment en/of angst oproepen bij het kind en richten de aandacht van het kind niet op de gevoelens en gedachten van het slachtoffer, maar op de negatieve gevolgen van de overtreding voor het kind zelf (straf of liefdesonthouding). Deze technieken zullen dan ook eerder een egoïstische dan een prosociale houding bij het kind versterken.

Hoewel Hoffmans verklaring plausibel lijkt en er in onderzoek van Hoffman zelf ook enige empirische ondersteuning gevonden is voor de gepostuleerde relaties (zie voor een overzicht Janssens & Gerris, 1984; Gerris & Janssens, 1984), stellen Radke-Yarrow, Zahn-Waxler en Chapman (1983) in hun overzichtstudie dat er zowel positieve, negatieve als geen relaties zijn gevonden tussen machtsuitoefening, liefdesonthouding en inductie enerzijds en prosociale ontwikkeling anderzijds. Een verklaring hiervoor kan gelegen zijn in het feit dat de gevolgen van een bepaalde disciplineringstechniek afhankelijk kunnen zijn van het gebruik van deze disciplineringstechniek in het totale ouderlijk opvoedingspatroon. "If power assertion is linked negatively with children's prosocial behaviors, it is likely to be associated at the same time with negative attitudes on the part of the parent and, most likely, with punitive assertion.... In the context of a positive and responsive parent power assertion may relate quite differently to prosocial behavior by children. The influence of induction on prosocial behavior is similarly modified by rearing context, particularly affective contextual factors" (Radke-Yarrow, Zahn-Waxler & Chapman, 1983, pp. 509-510). Ook Staub (1979) stelt dat men beter de invloed van opvoedingspatronen op de ontwikkeling van het kind kan bestuderen dan de effecten van specifieke technieken binnen die patronen. We komen daarop nog terug.

1.2 *Ondersteuning en prosociale ontwikkeling*

In die opvoedingspatronen lijkt een affectieve, responsieve relatie tussen ouder en kind van belang. Door verschillende auteurs (Rollins & Thomas, 1979; Hoffman, 1970, 1975) is gewezen op de positieve invloed van ouderlijke ondersteuning op zich op de prosociale ontwikkeling van het kind. Rollins en Thomas (1979) verstaan onder ondersteuning die opvoedingsgedragingen waardoor het kind zich op zijn gemak voelt in aanwezigheid van de ouder en waardoor bij het kind het idee post vat dat het als persoon door de ouder wordt geaccepteerd. De volgende gedragingen zijn daarvan voorbeelden: bemoedigen, prijzen, helpen, samenwerken en affectie tonen. Volgens Hoffman (1963) kan ondersteuning ertoe leiden dat het kind niet gefrustreerd wordt door eigen onvervulde behoeften die het kind blind maken voor de behoeften van anderen. Ondersteuning bevordert de emotionele zekerheid van het kind; het behoeft zich minder bezig te houden met de eigen belangen en krijgt meer oog voor die van anderen. Ook door Radke-Yarrow, Zahn-Waxler en Chapman (1983) wordt gewezen op een mogelijk mediërende rol van empathie tussen ouderlijke ondersteuning en de prosociale ontwikkeling van het kind. Ondersteuning door de ouder geeft het kind de ruimte met gevoelens en belangen van anderen rekening te houden en zich derhalve in te leven in en mee te leven met anderen. Toch concluderen zij in hun overzichtstudie dat de gevonden relaties tussen prosociale ontwikkeling en ondersteuning niet consistent in dezelfde richting wijzen; in sommige studies zijn wel, in andere zijn geen

relaties gevonden tussen ondersteuning en prosociale ontwikkeling. De auteurs maken daarbij als kanttekening, dat het in de meeste studies gaat om algemeen ondersteunend gedrag. Volgens hen is het van belang om niet in zijn algemeenheid naar ondersteunend gedrag te kijken, maar naar ondersteunend gedrag in situaties die voor het kind van direct belang zijn. Zij wijzen in dat verband op situaties waarin het kind in nood verkeert of leed ervaart. Ondersteuning krijgt in deze benadering de betekenis van responsiviteit, het gevoelig zijn voor de behoeften en gevoelens van het kind en daarop prompt en adequaat reageren. Ook Staub (1979) benadrukt het specifieke belang van ondersteuning in de vorm van responsiviteit. Als ouders responsief reageren, krijgt het kind het gevoel dat het controle heeft over zijn omgeving. Bovendien krijgt het kind het idee dat het vertrouwen kan hebben in die omgeving. Die gevoelens van vertrouwen en controle kunnen volgens Staub (1979) een positieve oriëntatie op andere mensen tot gevolg hebben.

Staub (1979) benadrukt daarbij dat dit responsief gedrag ook als zodanig ervaren moet worden door het kind. Ook uit onderzoek van Hoffman (1970) blijkt dat niet zozeer inductief gedrag zoals gerapporteerd door de ouders, maar inductief gedrag zoals beleefd door het kind samenhangt met de prosociale ontwikkeling van het kind. Volgens interactionistische verklaringsmodellen is de subjectieve (gedrags)omgeving zoals die door het kind wordt waargenomen en geïnterpreteerd duidelijker gerelateerd aan de individuele ontwikkeling dan de meer afstandelijk gedefinieerde en door volwassenen gerapporteerde omgeving (Magnusson & Allen, 1983). Het gaat ons derhalve om de opvoedingsdimensies zoals die door het kind worden beleefd. Uit eerder onderzoek is gebleken dat er weliswaar een duidelijke overeenkomst is tussen het opvoedingsgedrag zoals het wordt gerapporteerd door de ouders en het opvoedingsgedrag bezien met de ogen van het kind, toch zijn de gevonden correlaties ($r = .29$ voor vaders, $r = .38$ voor moeders) niet dusdanig hoog dat onderzoek van opvoedingsdimensies vanuit de waarneming van het kind niet zinvol zou zijn (cf. Gerris, Janssens & Janssen, 1988).

1.3 *Beroep doen op de verantwoordelijkheid van het kind en prosociale ontwikkeling*

In zijn latere publikaties heeft Hoffman (1976, 1982) niet alleen gewezen op het belang van inductieve disciplinering en ondersteunend gedrag van ouders voor de prosociale ontwikkeling van het kind, maar tevens op de betekenis van gedrag waarbij het kind wordt gewezen op zijn verantwoordelijkheid en wordt aangespoord tot rijp en aan zijn leeftijd aangepast gedrag. Ook door andere auteurs (Baumrind, 1971; Maccoby & Martin, 1983) wordt een belangrijke invloed van dat soort gedrag op de prosociale ontwikkeling van het kind verondersteld. Zij vatten dat gedrag veelal samen onder het concept 'demandingness'. Hieronder wordt gedrag van ouders verstaan waarbij een appèl wordt gedaan op het vermogen bij het kind tot zelfstandigheid, verantwoordelijkheid of gedrag waarbij rijp gedrag geëist wordt dat passend is bij de leeftijd van het kind of gedrag waarbij de ouder gericht is op de autonomie van het kind als persoon en waarbij de ouder bovendien toeziet op naleving van de gestelde eisen. Een goed, kort en bondig Nederlandstalig equivalent voor het concept 'demandingness' ontbreekt. In het volgende spreken we van een beroep doen op de verantwoordelijkheid van het kind. Ondanks de theoretische nadruk die er de laatste jaren op dit concept wordt gelegd, zijn er nauwelijks studies waarin dit begrip gerelateerd is aan prosociale ontwikkeling. In het overzichtsartikel van Maccoby en Martin (1983) worden alleen studies gerapporteerd waarin de invloed van 'demandingness' is bestudeerd in samenhang met ondersteuning. Zo vonden Yarrow, Waxler en Scott (1971) dat beide gedragingen nodig zijn. Ook concluderen Maccoby en Martin (1983) dat kinderen van ouders die ondersteunend zijn en tegelijkertijd eisen durven stellen aan de verantwoordelijkheid van hun kind, meer competent en meer onafhankelijk zijn en zich meer sociaal verantwoordelijk voelen. Eisenberg, Lennon en Roth (1983) stellen vast dat het stimuleren van autonomie van belang is voor de ontwikkeling van prosociaal gedrag, en Staub (1979) concludeert op basis van Baumrinds studies dat het stimuleren van verantwoordelijkheid en onafhankelijkheid belangrijk zijn voor het tot stand komen van altruïstisch gedrag.

Uit het overzichtsartikel van Maccoby en

Martin (1983) is duidelijk dat het moeilijk is om iets te zeggen over de effecten van 'demandingness' zonder daarbij te kijken naar de kwaliteit van ouderlijk ondersteunend gedrag. Uit een studie van Coopersmith (1967) blijkt bovendien dat gedrag waarbij ouders een beroep doen op de verantwoordelijkheid van hun kind gerelateerd is aan ouderlijke ondersteuning en controle. Ouders die hoge eisen stellen aan de competentie en gehoorzaamheid van hun kind en goed op naleving van die eisen toezien, zijn vaker inductief en hechten meer belang aan een democratisch patroon van beslissingen in het gezin.

Welke gevolgen hebben bovenstaande bevindingen voor onderzoek naar relaties tussen opvoedingsgedragingen en prosociale ontwikkeling? Op de eerste plaats hebben we geconcludeerd dat het belangrijk is om opvoedingsgedrag te meten zoals het door het kind wordt gepercipieerd. Op de tweede plaats is het van belang rekening te houden met de onderlinge afhankelijkheid van de verschillende opvoedingsgedragingen en de invloed ervan op prosociale ontwikkeling gelijktijdig te bestuderen. In ons onderzoek gaan we daarom eerst na of ondersteuning, de kwaliteit van disciplineren en het beroep doen op de verantwoordelijkheid van het kind aan elkaar gerelateerd zijn. Vervolgens analyseren we de relaties tussen deze aspecten van ouderlijk opvoedingsgedrag enerzijds en de prosociale ontwikkeling van het kind anderzijds. Ten slotte gaan we in op de mogelijk mediërende invloed van empathie tussen ouderlijk opvoedingsgedrag en de prosociale ontwikkeling van het kind.

1.4 *Empathie als interveniërende variabele*

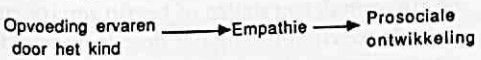
Tot nu toe hebben we aan die mediërende rol relatief weinig aandacht besteed. Er is evidentie voor de aanwezigheid van relaties tussen ouderlijk opvoedingsgedrag en empathie enerzijds en tussen empathie en prosociale ontwikkeling anderzijds. We hebben er al op gewezen dat Radke-Yarrow, Zahn-Waxler en Chapman (1983) een mediërende rol veronderstellen van empathie tussen ouderlijke ondersteuning en prosociale ontwikkeling. Ondersteuning door de ouder geeft het kind de ruimte met gevoelens en belangen van anderen rekening te houden en zich derhalve in te leven in en mee te leven met anderen. En volgens Hoffman en Saltzstein (1967) motiveert inductie het kind zich te oriënteren op de schade of het leed die het veroorzaakt heeft bij een ander. Het besef dat men de oorzaak is van dat leed of die schade kan empathische gevoelens doen ontstaan. Wellicht geldt iets soortgelijks voor gedrag waarbij de ouder een beroep doet op de verantwoordelijkheid van het kind. Als ouders eisen dat hun kind rekening houdt met de belangen van anderen, leren zij dat het kind mede verantwoordelijk is voor die anderen en dat besef kan bevorderen dat het kind zich inleeft in de gevoelens van die anderen. In tegenstelling tot relaties tussen empathie enerzijds en ouderlijk ondersteunend en disciplineringsgedrag anderzijds zijn relaties tussen empathie en gedrag van ouders waarbij een beroep wordt gedaan op de verantwoordelijkheid van het kind, nog nooit onderzocht.

Als empathie een interveniërende variabele is tussen opvoedingsgedrag en prosociale ontwikkeling, dan moeten er niet alleen relaties zijn tussen opvoedingsgedrag en empathie, maar ook tussen empathie en prosociale ontwikkeling. Niettemin concluderen Eisenberg en Miller (1987) in hun overzichtsartikel dat er slechts geringe samenhangen zijn tussen empathie en prosociale ontwikkeling. De sterkte van de samenhang hangt bovendien af van de soort maat die gebruikt wordt voor empathie. Positieve relaties doen zich vooral dan voor als er gebruik wordt gemaakt van een vragenlijst-methode om empathie te meten. In ons onderzoek hebben we dat gedaan door bij de kinderen de empathie-schaal van Bryant (1982) af te nemen.

Eisenberg en Miller noemen verschillende redenen waarom er vaak maar een geringe samenhang wordt gevonden tussen empathie en prosociale ontwikkeling. Een reden is gelegen in het feit dat er in de meeste studies maar van één maat voor prosociale ontwikkeling gebruik gemaakt wordt. In ons onderzoek hebben we verschillende maten betrokken. Op de eerste plaats proberen we prosociaal moreel redeneren te meten door de kinderen verhalen voor te leggen waarin het eigen belang conflicteert met dat van een ander en waarin het kind moet aangeven aan welk belang het de voorkeur geeft en waarom. Op de tweede plaats hebben we de leerkrachten van de in het onderzoek betrokken kinderen gevraagd hoe prosociaal het kind in de klas is. Hetzelfde is

door middel van een sociometrische procedure gevraagd aan de medeleerlingen.

Vatten we onze vraagstelling in een model samen, dan kunnen we stellen dat het in ons onderzoek gaat om een toets van het in Figuur 1 weergegeven model.



Figuur 1 *Verklaringsmodel voor prosociale ontwikkeling*

We zullen bij het toetsen van de adequaatheid van dit model nagaan of empathie inderdaad als een interveniërende variabele moet worden opgevat, dan wel als een afhankelijke variabele, die niet medieert tussen opvoedingsgedrag en prosociale ontwikkeling, maar die wel te verklaren is vanuit opvoedingsgedrag.

2 *Methode*

2.1 *Subjecten*

Er werden bij 120 Nederlandse gezinnen (beide ouders en een kind uit elk gezin) uit Nijmegen en omgeving data verzameld. De steekproef bestond uit 61 jongens en 59 meisjes. De leeftijd van de kinderen ligt tussen 9;5 en 12;11 jaar. Van de gezinnen komt 28% uit de lagere sociale klasse, 50% uit de middenklasse en 22% uit de hogere klasse (Van Westerlaak, Kropman & Collaris, 1975)

2.2 *Instrumenten*

Empathie is gemeten met behulp van Bryants empathie-schaal (Bryant, 1982). Deze schaal bestaat uit 22 items die in de vorm van een negenpuntsschaal aan het kind worden voorgelegd. Een voorbeeld van zo'n item is het volgende:

> Als ik een meisje zie huilen, moet ik ook bijna huilen.

Hoe meer empathische reacties het kind vertoont, des te hoger de score op deze schaal. De schaal heeft een betrouwbaarheid van .69 (alpha).

Prosociaal moreel redeneren is gemeten met behulp van vier aan Eisenberg, Lennon en Roth (1983) ontleende verhalen. In deze verhalen wordt het kind geconfronteerd met situaties waarin het moet kiezen tussen het eigen belang en dat van anderen. Aan de kinderen wordt gevraagd waaraan zij de voorkeur geven en waarom. Een voorbeeld van zo'n verhaal is het volgende:

> Meneer de Bruin liep een keer's avonds laat over straat. Opeens zag hij dat er een vrouw aan de overkant van de straat werd overvallen. Een ongewapende man probeerde haar tas af te pakken. Maar de vrouw hield de tas goed vast, en ze riep om hulp, terwijl ze met de dief aan het vechten was om haar tas. Meneer de Bruin keek snel om zich heen, maar hij zag niemand anders die zou kunnen helpen. De dief had hem nog niet gezien, dus hij kon er nog ongemerkt vandoor gaan. Hij wist niet goed wat hij moest doen. Want als hij zou helpen zou de dief hem misschien pakken.

De antwoorden van de kinderen zijn op audiorecorder opgenomen en daarna uitgeschreven en gecodeerd (mate van intersubjectieve overeenstemming = .83). Per verhaal werd genoteerd of het kind egoïstisch redeneerde en/of zich liet leiden door de belangen van anderen in de situatie en/of het kind redeneerde vanuit een empathische oriëntatie of vanuit geïnternaliseerde normen of gevoelens. Per kind werd één score berekend: in hoeveel verhalen redeneert het kind vanuit het belang van anderen of vanuit een empathische oriëntatie of vanuit geïnternaliseerde gevoelens of normen zonder ook maar één egoïstisch motief te noemen. De scores die derhalve lopen van nul tot en met vier, duiden op de mate waarin het kind redeneert vanuit het belang van anderen en niet vanuit eigenbelang.

Prosociaal gedrag werd vastgesteld door de leerkrachten van de 120 kinderen te vragen de Prosocial Behaviour Questionnaire van Weir en Duveen (1981) in te vullen. Deze bestaat uit 20 zespuntsitems. Elk item heeft betrekking op prosociaal gedrag van het kind in de klas. Een voorbeeld-item is het volgende:

> Biedt hulp aan andere kinderen die moeite hebben met een bepaalde opdracht in de klas.

De betrouwbaarheid van de schaal is .90 (alpha). Hoe hoger de score, hoe meer prosociaal gedrag het kind vertoont in de klas volgens de leerkracht.

Voor een tweede meting van prosociaal gedrag in de klas zijn de klasgenoten als informatiebron gebruikt. Aan de klasgenoten van de 120 kinderen werd gevraagd drie kinderen in hun klas te noemen die zij het meest mochten en drie die het meest anderen hielpen in de

klas. Voor elk van de 120 kinderen is vastgesteld hoe vaak ze op elke vraag afzonderlijk door klasgenoten genoemd zijn. Deze aantallen zijn vanwege verschillen in klassegrootte gedeeld door het aantal leerlingen in de klas. De scores geven derhalve de mate aan waarin het kind prosociaal gedrag vertoont in de ogen van de leerlingen.

Om de mate van ouderlijke *ondersteuning*, de mate waarin ouders een *beroep doen op de verantwoordelijkheid* van het kind en de kwaliteit van het *disciplineringsgedrag* van ouders te meten is de volgende procedure gevolgd. Aan de kinderen zijn acht hypothetische situaties voorgelegd waarin een kind een regel of norm overtreedt. Een voorbeeld van zo'n verhaal is het volgende:

> Je bent thuis samen met je vriend Kees aan het spelen met LEGO. Je wilt eigenlijk een paar stukjes LEGO hebben die Kees heeft. Maar Kees geeft die stukjes niet af, terwijl hij ze eigenlijk niet nodig heeft. Je wordt nijdig, loopt op Kees af en slaat hem tegen zijn hoofd. Kees begint te huilen.

De kinderen werd gevraagd wat hun vader en wat hun moeder zouden doen of zeggen als het kind zelf zo'n overtreding zou begaan. Alle antwoorden werden op audiorecorder opgenomen en uitgeschreven. Per situatie en per ouder werd nagegaan of een kind aangaf dat de ouder machtsuitoefenend, inductief, ondersteunend zou reageren en/of een beroep zou doen op de verantwoordelijkheid van het kind. Een antwoord of een deel daarvan werd gecodeerd als *machtsuitoefenend* als een ouder het kind zou straffen of daarmee zou dreigen of als de ouder louter verbaal geweld zou gebruiken. Een antwoord werd als *inductief* gecodeerd als een ouder zou wijzen op de gevolgen van de overtreding voor een ander of als die ouder het kind zou vragen zich eens voor te stellen hoe het zou zijn als het zelf slachtoffer was in zo'n situatie of als de ouder redenen gaf aan het kind zich in de toekomst anders te gedragen. Ouderlijk gedrag is als *ondersteunend* gecodeerd als de ouder het kind op zijn gemak zou stellen of begrip zou tonen voor de overtreding, omdat deze is uitgelokt door een ander. Ten slotte is per situatie gecodeerd of de ouder zou *wijzen op de verantwoordelijkheid van het kind* of zou aangeven welk gedrag hij/zij zou verwachten in de toekomst. Voor elke ouder zijn vervolgens vier scores berekend: in hoeveel situaties wordt het gedrag van de ouder door het kind als ondersteunend ervaren, in hoeveel als machtsuitoefenend, in hoeveel als inductief en in hoeveel situaties zou de ouder een beroep doen op de verantwoordelijkheid van het kind. Er is geen score voor liefdesonthouding berekend omdat dat gedrag nauwelijks genoemd werd door de kinderen.

3 *Resultaten*

In Tabel 1 zijn de relaties weergegeven tussen de verschillende opvoedingsgedragingen van moeders en vaders zoals ze door het kind worden ervaren. Tot een aparte weergave voor vaders en moeders is besloten omdat het per ouderpaar gaat om hetzelfde kind.

Uit Tabel 1 blijkt dat er een duidelijk onderling verband is tussen de onderscheiden opvoedingsgedragingen. Ondersteuning van het kind door de moeder hangt positief samen met inductie ($r = .24$) en negatief met machtsuitoefening ($r = -.19$). Verder blijkt dat moeders die het kind ondersteunen en inductief reageren op overtredingen, ook relatief vaker

Tabel 1 *Relaties tussen opvoedingsgedragingen van moeder (M) en vader (V) zoals ervaren door het kind (* = $p < 0.05$)*

	Ondersteuning		Inductie		Machts-uitoefening	
	M	V	M	V	M	V
Inductie	.24*	.11				
Machtsuitoefening	-.19*	-.27*	-.53*	-.54*		
Beroep doen op eigen verantwoordelijk	.16*	.08	.42*	.25*	-.50*	-.57*

Tabel 2 *Relaties tussen maten voor prosociale ontwikkeling en empathie* (* = $p<0.05$)

	Prosociaal gedrag volgens leerkracht	Prosociaal redeneren	Genoemd als meest geliefd	Genoemd als iemand die veel helpt
Prosociaal redeneren	0.25*			
Genoemd als meest geliefd	0.23*	0.21*		
Genoemd als iemand die veel helpt	0.39*	0.03	0.38*	
Empathie	0.16	0.18*	0.17*	0.10

een beroep doen op de eigen verantwoordelijkheid van het kind (resp. r = .16 en r = .42). In de ogen van het kind is dit beroep doen op de eigen verantwoordelijkheid door de moeder moeilijk verenigbaar (r = -.50) met machtsuitoefening. Eenzelfde conclusie geldt voor de samenhang tussen inductieve en machtsuitoefenende disciplinering (r = -.53)

De samenhangen tussen de opvoedingsgedragingen van de vaders zoals waargenomen door het kind zijn eveneens weergegeven in Tabel 1. Er is eenzelfde patroon van onderlinge samenhangen als bij de moeders: positieve relaties tussen ondersteuning, inductieve disciplinering en een beroep doen op de verantwoordelijkheid van het kind, waarbij echter alleen de relatie tussen inductie en het beroep doen op de eigen verantwoordelijkheid statistisch significant is (r = .25). De relaties tussen machtsuitoefening enerzijds en inductie en een beroep doen op de eigen verantwoordelijkheid anderzijds zijn sterk negatief (resp. r = -.54 en r = -.57).

Voordat we ingaan op relaties tussen de verschillende opvoedingsgedragingen en de prosociale ontwikkeling van het kind, bestuderen we eerst de relaties tussen de verschillende indicatoren voor deze ontwikkeling onderling en de relaties ervan met empathie. Deze samenhangen zijn weergegeven in Tabel 2.

Uit Tabel 2 blijkt dat empathie zwak, zij het significant samenhangt met prosociaal redeneren vanuit het belang van anderen (r = .18) en het geliefd zijn bij klasgenoten (r = .17), terwijl de correlaties van empathie met meer schoolse prosociale ontwikkelingsaspecten (zoals prosociaal gedrag volgens de leerkracht en het helpen van medeleerlingen) niet significant zijn. Opvallend in Tabel 2 is op de tweede plaats dat de school-gerelateerde prosociale ontwikkelingsaspecten hoog onderling correleren; kinderen die door hun leerkracht als prosociaal worden gezien, zijn vaker geliefd bij hun leeftijdgenoten (r = .23) en zij worden ook vaker door hun medeleerlingen gezien als iemand die anderen helpt (r = .39). Kinderen die vaak andere kinderen helpen zijn ook meer geliefd bij medeleerlingen (r = .38). Dit zou mogelijk kunnen duiden op domeinspecificiteit in de betreffende variabelen. Een derde bevinding betreft de positieve samenhang tussen prosociaal gedrag en prosociaal redeneren; kinderen die volgens de leerkracht prosociaal zijn en bij hun klasgenoten geliefd zijn, redeneren vaker vanuit het belang van anderen en minder vanuit eigenbelang (resp. r = .25 en r = .21). Een laatste bevinding uit Tabel 2 is de opvallend lage niet significante correlatie tusssen a) het helpen van klasgenoten en prosociaal redeneren en b) het helpen van klasgenoten en empathie. Mogelijk houdt deze bevinding ook verband met het domein-gebonden karakter van de 'help'-variabele, namelijk school- en klasgerelateerde taken, en niet persoon- en relatiegerichte situaties. Uit bovenstaande analyse van Tabel 2 kan men concluderen dat de hier gebruikte set van variabelen voor prosociale ontwikkeling uiteenvalt in twee groepen:

1. Prosociale gedragsaspecten in schoolse taak- en omgangssituaties. Deze gedragsaspecten betreffen vooral de extern voor leerkracht en medeleerlingen waarneembare, opvallende gedragingen.
2. Meer aan de persoon gebonden aspecten van prosociale ontwikkeling (als persoon

Tabel 3 *Relaties tussen opvoeding door moeder (M) en vader (V) zoals ervaren door het kind en prosociale ontwikkeling en empathie* (* = $p<0.05$)

	Ondersteuning		Inductie		Machts-uitoefening		Beroep doen op verantwoordelijkheid	
	M	V	M	V	M	V	M	V
Prosociaal gedrag volgens leerkracht	.27*	.05	-.05	-.11	-.05	-.02	.03	.03
Prosociaal redeneren	.22*	.05	.18*	.04	-.26*	-.07	.18*	.17*
Genoemd als meest geliefd	.04	.01	.22*	.26*	-.21*	-.22*	.19*	.07
Genoemd als iemand die vaak helpt	-.05	-.06	.07	.13	-.04	-.06	.09	.00
Empathie	.12	.04	.22*	.31*	-.28*	-.21*	.34*	.23*

geliefd zijn en rekening houden met het belang van een ander).

De samenhangen tussen de verschillende opvoedingsgedragingen van vader en moeder enerzijds en de mate van prosociale ontwikkeling en empathie van het kind anderzijds zijn weergegeven in Tabel 3.

Uit Tabel 3 blijkt dat ondersteuning ervaren bij de moeder positief samenhangt met prosociaal gedrag in de klas (r = .27) en met prosociaal redeneren (r = .22). Empathie hangt niet significant samen met ondersteuning. Wel is er een positieve samenhang tussen empathie en inductie (r = .22) en tussen empathie en het beroep doen op de eigen verantwoordelijkheid van het kind (r = .34). Empathie hangt daarentegen negatief samen met machtsuitoefening (r = -.28). In Tabel 3 zien we verder dat inductie significant samenhangt met prosociaal redeneren (r = .18) en het genoemd worden als meest geliefd door klasgenoten (r = .22). Ook het beroep doen op de eigen verantwoordelijkheid van het kind hangt positief samen met prosociaal redeneren (r = .18) en het geliefd zijn bij klasgenoten (r = .19). Machtsuitoefening hangt daarentegen duidelijk negatief samen met prosociaal redeneren (r = -.26) en het geliefd zijn bij klasgenoten (r = -.21). Op grond van deze bevindingen in Tabel 3 kan men concluderen dat het empathisch vermogen van het kind, zijn vermogen om het belang van de ander te laten prevaleren boven het eigen belang en de populariteit bij klasgenoten worden bevorderd door inductief gedrag van de moeder en gedrag waarbij zij een beroep doet op de eigen verantwoordelijkheid van het kind.

Wanneer we de relaties bestuderen tussen empathie en prosociale ontwikkeling enerzijds en het opvoedingsgedrag van vader anderzijds, zien we een vergelijkbaar patroon van samenhangen (zie Tabel 3). Ook hier geldt een sterke samenhang tussen inductie en empathie (r = .31) en tussen empathie en het een beroep doen op de eigen verantwoordelijkheid van het kind (r = .23). Populariteit bij leeftijdgenoten hangt positief samen met inductie (r = .26) en prosociaal redeneren correleert positief met het beroep doen op de eigen verantwoordelijkheid van het kind (r = .17). Verder blijkt ook bij vaders een consistent negatieve samenhang tussen machtsuitoefening enerzijds en empathie (r = -.21) en populariteit bij klasgenootjes (r = -.22) anderzijds. In tegenstelling tot ondersteuning ervaren van de kant van moeder vertoont ondersteuning door vader geen samenhang met empathie of met prosociaal redeneren. Een ander verschil tussen vaders en moeders is dat voor vaders alleen de mate waarin zij een beroep doen op de verantwoordelijkheid van hun kind samenhangt met prosociaal redeneren door hun kind, terwijl dat redeneren samenhangt met alle vier de onderscheiden opvoedingsgedragingen van moeder.

Zowel bij de moeders als de vaders (zie Tabel 3) valt de consistent significant positieve samenhang op tussen inductie en een beroep doen op de eigen verantwoordelijkheid van het kind enerzijds en anderzijds empathie en de persoonsgebonden aspecten van prosociale ontwikkeling (prosociaal redeneren en geliefd zijn bij leeftijdgenoten), en de consistent negatieve samenhang tusen machtsuitoefening

Tabel 4 *Latente factoren in de Lisrel-analyse van moeders (M) en vaders (V) en daarop ladende variabelen (inclusief de hoogte van die ladingen)*

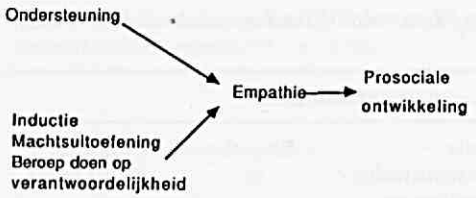
	Latente Factoren							
	Ondersteuning		Inductie Machtsuitoefening Beroep doen op veranwoordelijkheid		Empathie		Prosociale ontwikkeling	
	M	V	M	V	M	V	M	V
Variabelen								
- Ondersteuning	1.00	1.00						
- Inductie			0.67	0.54				
- Machtsuitoefening			-0.77	-1.00				
- Beroep doen op verantwoordelijkheid			0.66	0.57				
- Empathie					1.00	1.00		
- Prosociaal gedrag volgens leerkracht							0.67	0.58
- Prosociaal redeneren							0.34	0.28
- Genoemd als meest geliefd							0.48	0.54
- Genoemd als iemand die vaak helpt							0.54	0.62

en empathie en dezelfde persoonsgebonden aspecten van prosociale ontwikkeling. De schoolgebonden aspecten van prosociale ontwikkeling (door de leerkracht en medeleerlingen gezien worden als iemand die helpt) blijken nauwelijks met de onderscheiden opvoedingsgedragingen samen te hangen. Alleen blijkt dat naarmate moeders meer ondersteuning geven aan hun kind, het oordeel van de leerkracht over het prosociale gedrag van het kind in de schoolse situatie positiever is.

Uit de data hebben we geconcludeerd dat inductie, machtsuitoefening en beroep doen op de eigen verantwoordelijkheid van het kind aan elkaar gerelateerde concepten zijn. Deze concepten moeten daarom tegelijkertijd bestudeerd en geanalyseerd worden als het gaat om de invloed ervan op empathie en prosociale ontwikkeling. Dat is mogelijk met behulp van Lisrel-analyse. Ten behoeve van deze analyse construeerden we twee latente opvoedingsfactoren. De eerste is een factor met positieve ladingen van inductie en een beroep doen op de verantwoordelijkheid van het kind en een negatieve lading van machtsuitoefening. Uit Tabel 1 hebben we geconcludeerd dat deze gedragingen onderling hoog correleren; om die reden hebben we ze laten laden op één factor. Ouderlijke ondersteuning hangt weliswaar enigszins samen met de genoemde opvoedingsgedragingen, maar de samenhangen zijn dermate laag dat we een tweede latente opvoedingsfactor in het model hebben opgenomen met daarop alleen een lading van de variabele 'ondersteuning'. In Tabel 4 is voor moeders en voor vaders aangegeven hoe hoog elk opvoedingsgedrag laadt op de twee latente opvoedingsfactoren.

Op grond van de resultaten in Tabel 3 veronderstellen we dat met name de eerste latente factor een positieve invloed heeft op een latente empathie-factor, die samenvalt met de variabele 'empathie'. Empathie blijkt immers positief samen te hangen met inductie en een beroep doen op de eigen verantwoordelijkheid van het kind en negatief met machtsuitoefening. Hoewel ook op theoretische gronden een invloed van ondersteuning op empathie zou mogen worden verwacht, lijkt deze invloed op basis van Tabel 3 minder waarschijnlijk. Ondersteuning correleert immers nauwelijks met empathie.

Ook voor 'prosociale ontwikkeling' is gelet op de onderling hoge correlaties tussen de indicatoren voor dit concept één latente factor

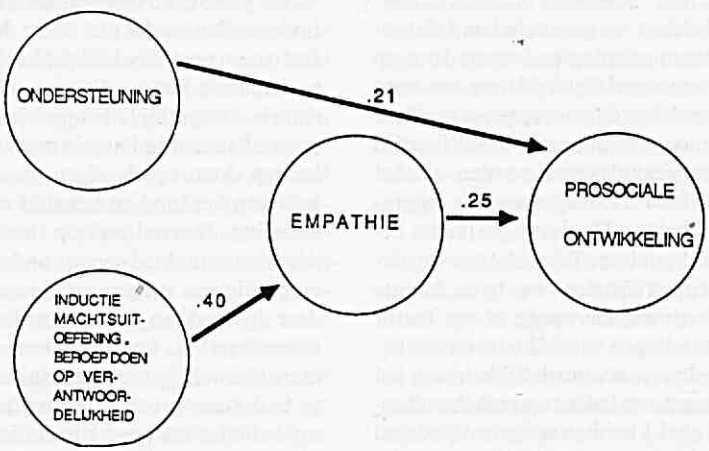


Figuur 2 *Te toetsen Lisrel-model ter verklaring van prosociale ontwikkeling vanuit de opvoeding zoals ervaren door het kind met empathie als interveniërende variabele*

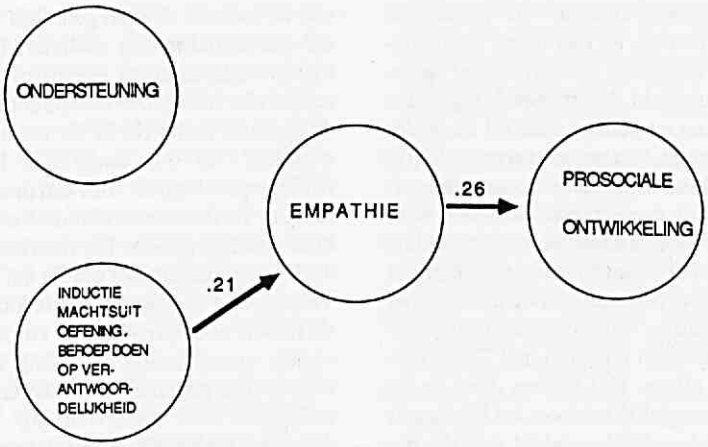
geconstrueerd, met daarop vier positief ladende variabelen: prosociaal gedrag in de klas, de twee oordelen van medeleerlingen en prosociaal redeneren. De ladingen van de vier prosociale variabelen op de ene latente factor 'prosociale ontwikkeling' zijn in Tabel 4 weergegeven. Op basis van de resultaten in Tabel 2 verwachten we een positieve invloed van de empathie-factor op de factor voor prosociale ontwikkeling.

Het te toetsen Lisrel-model is weergeven in Figuur 2. In Figuur 3 en 4 zijn de resultaten van de Lisrel-analyse voor resp. moeders en vaders weergegeven. Zowel het model voor vaders als voor moeders geeft een bevredigende passing (zie de chi-kwadraat-gegevens bij de Figuren 3 en 4). Ook de goodness-of-fit index (GFI) en de adjusted goodness-of-fit index (AGFI) zijn hoog; resp. .92 en .87 voor moeders en .92 en .88 voor vaders (zie voor criteria Lapsley & Quintana, 1989; Tanaka, 1987). Het is echter zinvol om voor de beoordeling van het model niet alleen op de totale passing af te gaan, maar tevens op de stabiliteit van de afzonderlijke coëfficiënten in het model. Alle geschatte ladingen (zie Tabel 4) en alle geschatte regressie-coëfficiënten – de waarden van de regressie-coëfficiënten zijn langs de pijlen in de Figuren 3 en 4 weergegeven – wijken statistisch significant af van nul. Dat duidt op redelijk stabiele coëfficiënten.

Uit de Figuren 3 en 4 blijkt op de eerste plaats dat de invloed van ouderlijke ondersteuning op prosociale onwikkeling gering is. Alleen voor moeders geldt dat er een directe invloed is van deze opvoedingsfactor op de prosociale ontwikkeling van het kind. Op de tweede plaats is er een bescheiden positieve, maar wel statistisch significante invloed van empathie op prosociale ontwikkeling. Daarnaast blijkt dat empathie afhankelijk is van een opvoedingspatroon dat gekenmerkt wordt door inductie, een beroep doen op de eigen verantwoordelijkheid van het kind en weinig machtsuitoefening. Hieruit concluderen we dat empathie als een interveniërende variabele tussen dat opvoedingspatroon en prosociale ontwikkeling kan worden beschouwd. Om deze bevinding verder te onderbouwen zijn er alternatieve modellen getoetst. In een eerste model is empathie niet als interveniërende, maar als afhankelijke variabele opgenomen naast de afhankelijke variable 'prosociale ontwikkeling'. Dat model levert een slechtere passing op dan de modellen in Figuur 3 en 4 (chi-kwadraat = 40.83; df = 27; p = 0.043; GFI = -.92; AGFI = .86 voor moeders en chi-kwadraat = 42.54; df = 29; p = 0.050; GFI = .92; AGFI = .87 voor vaders). Nog slechter is de passing van een model



Figuur 3 *Lisrel-model voor moeders (chi-kwadraat = 37.07; df = 26; p = 0.074)*



Figuur 4 *Lisrel-model voor vaders (chi-kwadraat = 38.45; df = 28; p = 0.090)*

waarin empathie niet als een zelfstandige latente factor is opgenomen, maar is opgevat als een indicator voor de latente factor 'prosociale ontwikkeling' (chi-kwadraat = 51.70; df = 26; p = 0.002; GFI = .90; AGFI = .83 voor moeders en chi-kwadraat = 51.37; df = 27; p = 0.003; GFI = .89; AGFI = .82 voor vaders).

4 *Discussie*

Uit de onderlinge samenhang tussen de onderscheiden opvoedingsgedragingen is de conclusie getrokken dat ondersteuning een aparte opvoedingsdimensie vormt die relatief onafhankelijk is van een tweede opvoedingsdimensie die gekenmerkt wordt door veel inductie, een beroep doen op de eigen verantwoordelijkheid van het kind en weinig gebruik van machtsuitoefening. Deze laatste opvoedingsdimensie zouden we willen omschrijven als een autoritatieve factor. Daarbij hanteren we deze term niet in de betekenis die er door Maccoby en Martin (1983) aan is gegeven, maar in de betekenis die Baumrind (1971) er oorspronkelijk aan heeft gegeven. In het model van Maccoby en Martin (1983) wordt het autoritatieve opvoedingspatroon gekenmerkt door een beroep doen op de eigen verantwoordelijkheid van het kind én ondersteuning. Volgens Maccoby en Martin (1983) zijn autoritatieve ouders ouders die én ondersteunend zijn én vaak een beroep doen op de eigen verantwoordelijkheid van hun kind. Uit ons onderzoek moeten we concluderen dat ondersteuning relatief onafhankelijk is van het autoritatieve opvoedingspatroon. Als Baumrind (1971) het heeft over een autoritatief opvoedingspatroon heeft ze het niet over ouderlijke ondersteuning. Autoritatief opvoedingsgedrag is volgens haar gedrag waarbij ouders een beroep doen op de zelfstandigheid en verantwoordelijkheid van het kind en op naleving van redelijke eisen toezien. Uit ons onderzoek blijkt dat die ouders meer inductief en minder machtsuitoefenend reageren op hun kind. Beschrijven we een autoritatieve ouder als een ouder die vaak een beroep doet op de eigen verantwoordelijkheid van het kind, vaak inductief en weinig machtsuitoefenend reageert, dan kunnen we een 'autoritaire' ouder beschrijven als een ouder die vaak machtsuitoefenend en weinig inductief reageert, en zelden een beroep doet op de eigen verantwoordelijkheid van het kind. De autoritaire ouder beperkt de gedragsalternatieven van het kind, accepteert geen afwijkingen van opgelegde regels en geeft geen verantwoordelijkheid aan het kind. De nadruk ligt meer op wat het kind fout doet, terwijl de autoritatieve ouder vooral geïnteresseerd is in rijp, aan de leeftijd aangepast gedrag van het kind. Deze dimensie van autoritatief versus autoritair gedrag is relatief onafhankelijk van de ondersteuningsdimensie. In de Lisrel-analyse wordt de correlatie tussen deze dimensies voor moeders geschat op .28 en voor vaders op .27.

Wat de vraagstelling naar de samenhang tussen opvoedingsgedragingen en prosociale ontwikkeling betreft, en met name de mediërende rol van empathie daarbij wijzen de resultaten eenduidig in dezelfde richting: voor inductief gedrag en gedrag waarbij de ouder een beroep doet op de eigen verantwoordelijkheid van het kind functioneert empathie als een duidelijk interveniërende variabele naar de persoonsgebonden aspecten van prosociale ontwikkeling. Uit de analyses is gebleken dat in een autoritatief opvoedingspatroon met het oog op het criterium van prosociale ontwikkeling niet kan worden volstaan met inductieve disciplinering alleen. Het beroep doen op de eigen verantwoordelijkheid van het kind geeft een zodanig inhoudelijk gewicht aan dit patroon dat alleen in die combinatie een positieve uitwerking naar empathie en prosociale ontwikkeling mogelijk is. Dit impliceert voor de alledaagse opvoeding dat het zich inleven en rekening houden met de gevoelens van anderen in de ervaring van kinderen pas goed tot hun recht komen, als ouders (moeders en vaders) uitleggen op welke wijze het kind verantwoordelijk is voor zijn slachtoffer (inductie) en eisen stellen ten aanzien van verantwoordelijk gedrag. Om een bijdrage te kunnen leveren aan de vorming van meevoelende mensen is het dus vooral nodig dat ouders een duidelijk appèl doen op de eigen verantwoordelijkheid van het kind voor de gevolgen van zijn gedrag en laten zien op welke wijze het kind zelf (mede)verantwoordelijk is voor het verdriet van de ander.

De uitkomsten van de Lisrel-analyse betekenen overigens niet dat alleen empathie zou mediëren tussen het beschreven opvoedingspatroon en prosociale ontwikkeling. De gevonden regressie-coëfficiënten zijn relatief laag. Er zijn derhalve ook andere factoren dan empathie die ertoe leiden dat een autoritatief opvoedingspatroon een positieve invloed heeft op de prosociale ontwikkeling van het kind. Dat geldt zeker voor de invloed van ondersteunend opvoedingsgedrag. Uit het Lisrel-model blijkt dat deze invloed van ondersteuning van de kant van de moeder niet wordt gemediëerd door empathie. De invloed van ondersteuning moet derhalve vanuit andere mechanismen worden verklaard. Een eerste interpretatie is geformuleerd in termen van de sociale leertheorie. Wellicht functioneert een ondersteunende moeder als model voor haar kind. Het kind neemt het gedrag van de moeder over en gedraagt zich prosociaal ten opzichte van anderen. De correlatie tussen ondersteuning en prosociaal gedrag is echter niet hoog. Dat is wellicht te wijten aan de geringe variantie in de variabele 'ondersteuning'. In ons onderzoek leenden zich slechts twee situaties voor een ondersteunende reactie. In deze twee situaties overtreedt het kind een regel nadat het daartoe is uitgelokt door een ander kind of omdat het verdrietig is. Vanwege de geringe variantie kan de correlatie tussen ondersteuning en empathie of prosociale ontwikkeling niet hoog worden. Een tweede interpretatie hangt samen met de mogelijkheid dat ondersteuning hier vooral doorwerkt naar schoolgerelateerde aspecten van prosociale ontwikkeling. In dit opzicht zou ondersteuning niet alleen een zekere modelwerking kunnen hebben voor aangepast sociaal gedrag maar ook kunnen werken als emotionele ondersteuning van het kind bij het cognitief verwerken van school of leerkracht gebonden frustraties van alledag. Kinderen die thuis een veilige uitlaatkap en opvang hebben voor deze dagelijkse situaties worden geholpen deze frustraties te verwerken waardoor ze niet meer kunnen 'ontaarden' in onaangepast gedrag in de klas.

Een belangrijke bevinding in ons onderzoek betreft ook de relatief lage, maar wel statistisch significante relatie tussen empathie en prosociale ontwikkeling. Het lijkt erop dat in dit onderzoek de conclusie van Eisenberg en Miller (1988) wordt bevestigd dat er slechts betrekkelijk lage correlaties gevonden zijn tussen empathie en prosociale ontwikkeling.

Ook de relatief lage, maar wel significante relatie tussen prosociaal redeneren en prosociaal gedrag is opvallend. In de literatuur treft men nauwelijks gegevens aan over deze relatie. Ook de verschillende indices voor prosociale ontwikkeling hangen positief met elkaar samen. Er is echter één meetinstrument voor prosociale ontwikkeling dat niet met de andere instrumenten voor dat begrip en niet met empathie samenhangt: het aan medeleerlingen vragen om drie leerlingen te noemen die het meest andere leerlingen helpen. We hebben daarvoor de volgende verklaring. Als we kinderen die laatste vraag stellen, denken ze waarschijnlijk vooral aan kinderen die anderen helpen met reken- of taalproblemen of

met andere meer cognitieve activiteiten. Zij noemen de 'slimme' kinderen uit de klas en niet de meest prosociale. Vragen we echter de drie kinderen te noemen die ze het meest aardig vinden in de klas, dan denken ze waarschijnlijk wel aan kinderen die aardig en eerlijk zijn voor anderen, die anderen op hun gemak stellen of die dingen delen met andere kinderen. Het zou kunnen zijn dat het antwoord op de vraag de drie meest geliefde kinderen te noemen een betere indicator is voor de prosociale ontwikkeling dan het antwoord op de vraag de drie kinderen te noemen die het meest anderen helpen.

Ten slotte willen we nogmaals wijzen op het belang om informatie over opvoedingsgedragingen van ouders te verzamelen bij de kinderen. Vaak wordt gesteld dat vraaggesprekken met ouders geen feitelijk gedrag meten. Strikt genomen geldt dat ook niet als men die vragen aan de kinderen stelt. Wel krijgt men op deze manier in elk geval informatie over de wijze waarop het kind het ouderlijk opvoedingsgedrag beleeft. Bovendien blijkt dat opvoedingsgedrag zoals door het kind ervaren samenhangt met empathie en prosociale ontwikkeling. Deze resultaten onderstrepen de aanbeveling van Staub (1979) om juist van deze ervaringen gebruik te maken.

Literatuur

Baumrind, D., Current patterns of parental authority. *Development psychology Monograph. 4*, 1971.

Bryant, B., An index of empathy for children and adolescents. *Child Development*, 1982, *53*, 413-425.

Coopersmith, S., *The antecedents of self-esteem.* San Francisco: Freeman, 1967.

Eisenberg, N., P. A. Miller, The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, 1987, *101*, 91-119.

Eisenberg, N., R. Lennon & K. Roth, Prosocial development: a longitudinal study. *Developmental Psychology*, 1983, *19*, 846-855.

Gerris, J. R. M. & J. M. A. M. Janssens, De invloed van inductie op morele internalisatie en altruïsme. *Tijdschrift voor Orthopedagogiek*, 1984 *23*, 458-474.

Gerris, J. R. M. & J. M. A. M. Janssens, Opvoedingsgedrag van ouders in overtredingssituaties: een onderzoek naar beïnvloedende factoren. *Pedagogische Studiën*, 1987, *64*, 295-307.

Gerris, J. R. M., J. M. A. M. Janssens & A. W. H. Janssen, Niveaus van pedagogisch perspectiefnemen van ouder en kind in disciplineringssituaties: Een empirisch verklaringsmodel. In: J. van Acker & J. R. M. Gerris (red.), *Gezinsrelaties onderzocht*. Amsterdam: Swets & Zeitlinger, 1988.

Hoffman, M. L., Parent discipline and the child's consideration for others, *Child Development*, 1963, *34*, 573-588.

Hoffman, M. L., Moral development. In: H. P. Mussen (Ed.), *Carmichael's Handbook of Child Psychology*. New York: Wiley, 1970.

Hoffman, M. L., Altruistic behavior and the parent-child relationship, *Journal of Personality and Social Psychology*, 1975, *11*, 937-943.

Hoffman, M. L., Empathy, roletaking, guilt and development of altruistic motives. In: Th. Lickona (Ed.), *Moral development and moral behavior.* New York: Holt, Rinehart & Winston, 1976.

Hoffman, M. L., Development of prosocial motivation; empathy and guilt. In: N. Eisenberg (Ed.), *The development of prosocial behavior.* New York: Academic Press, 1982.

Hoffman, M. L. & H. D. Saltzstein, Parent discipline and the child's moral development. *Journal of Personality and Social Psychology*, 1967, *5*, 45-57.

Janssens, J. M. A. M. & J. R. M. Gerris, Ouderlijke disciplinering en de ontwikkeling van morele internalisatie en altruïsme. *Tijdschrift voor Orthopedagogiek*, 1984, *23*, 386-399.

Janssens, J. M. A. M. & J. R. M. Gerris, Sociaal milieu en reacties van ouders op disciplineringssituaties; een empirisch verklaringsmodel. *Pedagogische Studiën*, 1988, *65*, 185-197.

Lapsley, D. K. & S. M. Quintana, Mental capacity and role taking; a structural equations approach. *Merrill Palmer Quarterly*, 1989, *35*, 143-163.

Maccoby, E. E., *Social development.* New York: Harcourt Brace Jovanovitch, 1980.

Maccoby, E. E. & J. A. Martin, Socialization in the context of the family: parent-child interaction. In: P. H. Mussen (Ed.), *Handbook of Child Psychology. Vol. IV. Socialization, personality and social development.* New York: Wiley, 1983.

Magnusson, D. K. & V. L. Allen, An interactional perspective for human development. In: D. Magnusson & V. L. Allen (Eds.), *Human development. An interactional perspective.* New York: Academic Press, 1983.

Radke-Yarrow, M., C. Zahn-Waxler & M. Chapman, Children's prosocial dispositions and behavior. In: P. H. Mussen (Ed.), *Handbook of child psychology, Vol. 4.* New York: Wiley, 1983.

Rollins, B.C. & D.L. Thomas, Parental support, power and control techniques in the socialization of children. In: W.R. Burr, R. Hill, F.I. Nye & I.L. Reiss (Eds.), *Contemporary theories about the family, Vol. 1.* London: Free Press, 1979.

Staub, E. *Positive social behavior and morality, Vol. 2.* New York: Academic Press, 1979.

Tanaka, J.S., "How big is big enough?", Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, 1987, *56*, 134-146.

Weir, K. & G. Duveen, Further development and validation of the Prosocial Behaviour Questionnaire for use by teachers. *Journal of child psychology and psychiatry*, 1981, *22*, 357-374.

Westerlaak, J. van, J. Kropman & J. Collaris, *Beroepenklapper.* Nijmegen: ITS, 1975.

Yarrow, M.R., C.Z. Waxler & P.M. Scott, Child effects on adult behavior. *Developmental Psychology*, 1971, *5*, 300-311.

Curricula vitae

J.M.A.M. Janssens (1947) is als universitair hoofddocent verbonden aan de vakgroep Algemene Pedagogiek van de Katholieke Universiteit Nijmegen. Zijn onderzoek heeft betrekking op relaties tussen opvoedingsoriëntaties en opvoedingsgedrag van ouders en de ontwikkeling van kinderen.

J.R.M. Gerris (1946) is als hoogleraar verbonden aan de vakgroep Algemene Pedagogiek van de Katholieke Universiteit Nijmegen. Zijn onderzoek heeft betrekking op opvoedingsprocessen in het gezin en andere primaire samenlevingsverbanden.

A.W.H. Janssen (1958) is als docent verbonden aan de afdeling Pedagogiek binnen de Educatieve Faculteit van de Hogeschool Gelderland. Hij verricht onderzoek naar sociale, morele en persoonlijkheidsontwikkeling van kinderen en de relatie daarvan met opvoeding in het gezin.

Adres: Vakgroep Algemene Pedagogiek Katholieke Universiteit Nijmegen, Erasmusplein 1, 6525 HT Nijmegen

Manuscript aanvaard 29-3-'90

Summary

Janssens, J.M.A.M., J.R.M. Gerris & A.W.H. Janssen. 'Child rearing, empathy and prosocial development.' *Pedagogische Studiën*, 1990, *67*, 403-416.

In this article we report the results of a project investigating relations between childrearing as perceived by the child, the child's empathy and prosocial development. Data were obtained from 120 Dutch families. The sample consisted of 61 boys and 59 girls between 9 and 12 years old. Our first conclusion is that in general perceived parental support, demandingness and induction are positively associated with the child's prosocial development and power assertion negatively. Secondly, we conclude that these childrearing practices are interrelated. Demandingness and induction are positively related and both have strong negative relations with power assertion. A third conclusion is that empathy has a positive influence on prosocial development. Moreover, empathy is dependent on a childrearing pattern that is characterized by induction and demandingness and in which parents use few power assertion. From this we conclude that empathy is a mediating link between childrearing and prosocial development.

Determinanten van TV-opvoeding: een replicatie-onderzoek

P. NIKKEN
Stichting Kinderpostzegels Nederland
J. E. VAN LIL
NOS, Afdeling Kijk- en Luisteronderzoek

Samenvatting

Via een telefonisch interview afgenomen bij een representatieve steekproef van 771 ouders met kinderen van 3 tot 18 jaar is onderzocht welke opvattingen ouders huldigen over de effecten van televisie op kinderen en hoe frequent zij tv-begeleidingsactiviteiten bij hun kinderen toepassen. Evenals in het Amerikaanse onderzoek dat we hebben gerepliceerd (Bybee, Robinson & Turow, 1982) laten tv-begeleidingsactiviteiten zich in drie soorten onderscheiden: restrictieve, evaluatieve en niet-gerichte begeleiding. Ouders zijn meer restrictief als ze televisie negatieve effecten toeschrijven, bij jongere kinderen en als er minder televisie-toestellen in huis zijn. Ouders begeleiden vaker evaluatief wanneer ze televisie positieve effecten toeschrijven en wanneer ze televisie negatieve effecten toeschrijven. Ouders begeleiden vaker niet-gericht als ze televisie positieve effecten toeschrijven, in kleinere gezinnen en als zij zelf meer televisie kijken. Moeders en hoger opgeleide ouders passen alle vormen van tv-begeleiding het meest toe.

1 *Inleiding*

1.1 *TV-opvoeding*

Recentelijk is er in de Nederlandse literatuur een sterk gestegen belangstelling aanwijsbaar voor de wijze waarop ouders hun kinderen bij het omgaan met televisie begeleiden (Vooijs, 1987; Van Lil & Vooijs, 1989; Nikken, 1989; Van der Voort, 1989). In dit verband wordt wel gesproken van 'tv-opvoeding', die door Bybee, Robinson en Turow (1982) in drie vormen wordt onderverdeeld: *restrictieve begeleiding* (controle uitoefenen op de tijd die kinderen aan televisie kijken besteden en op de programma's die zij op televisie zien), *niet-gerichte begeleiding* (samen met de kinderen meekijken naar televisie en daarover praten), en *evaluatieve begeleiding* (uitleg geven over programma-inhouden aan kinderen tijdens het kijken naar televisie).

Blijkens experimentele studies, waarvan Corder-Bolz (1982) en Vooijs (1987) een overzicht geven, kunnen ouders via verschillende tv-begeleidingsactiviteiten de leereffecten van televisie bij kinderen verhogen, alsook de houding van kinderen tegenover televisiebeelden beïnvloeden, en daarmee ook de uitwerking van die beelden op het gedrag. Niet-gerichte begeleidingsactiviteiten hebben een positieve uitwerking op wat kinderen van educatieve programma's kunnen leren (Ball & Bogatz, 1970), terwijl het toepassen van de evaluatieve begeleidingsactiviteiten angstreacties bij als griezelig ervaren programma's (Cantor & Wilson, 1984) en de aantrekkelijkheid en de invloed van geweldsbeelden op het agressieve gedrag van kinderen (Corder-Bolz, 1980) kunnen reduceren. Een restrictie van de tijd die kinderen aan televisie kijken besteden, kan bevorderen dat kinderen meer gaan lezen (Gadberry, 1974; De Meyer, Hendriks & Fauconnier, 1985).

Hoewel ouderlijke begeleidingsactiviteiten bij het televisie kijken effectief kunnen zijn, lijken weinig ouders van die mogelijkheid tot beïnvloeding gebruik te maken. Amerikaans onderzoek suggereert dat ouders nog het meest de niet-gerichte begeleiding toepassen en in mindere mate de restrictieve of evaluatieve begeleiding (Barcus, 1969; Bower, 1973; Mohr, 1979; Bybee et al., 1982; Palmer, Hockett & Dean, 1983; Messaris & Kerr, 1983).

Wat er in Nederlandse gezinnen aan tv-opvoeding wordt gedaan is niet vaak onderzocht. Slechts twee onderzoeken geven enig inzicht in de begeleidingsactiviteiten van ouders: een onderzoek van tien jaar geleden (Van Lil, 1978) en een recent onderzoek (Van Lil & Vooijs, 1989). Beide onderzoeken wijzen erop dat Nederlandse ouders vaker niet-gerichte begeleidingsactiviteiten toepassen dan restrictieve.

Comstock, Chaffee, Katzman, McCombs

en Roberts (1978) hebben de hypothese opgeworpen, dat de frequentie waarmee ouders tv-opvoedingsactiviteiten toepassen, voorspeld kan worden door de mate waarin ouders bezorgd zijn over de invloed van televisie op kinderen. In twee onderzoeken is feitelijk nagegaan of de opvattingen van ouders inderdaad van invloed zijn op het tv-opvoedingsgedrag. Bybee et al. (1982) verrichtten hun onderzoek bij een selecte groep ouders (200 massacommunicatie-wetenschappers met kinderen in de leeftijd van 3 tot 18 jaar). Bybee et al. (1982) verwachtten dat massacommunicatie-wetenschappers meer dan gewone ouders geneigd zouden zijn tot het uitoefenen van tv-begeleiding, hoewel dat na vergelijking met andere Amerikaanse studies verricht onder representatieve steekproeven, niet het geval leek te zijn. De ondervraagde wetenschappers hadden, blijkens factor-analyse, een genuanceerd oordeel over de rol van televisie; zij onderkenden tegelijkertijd zowel de positieve als de negatieve consequenties van televisie kijken voor kinderen. Bybee et al. (1982) spreken daarbij van pro-sociale en anti-sociale effecten. Conform de hypothese waren de wetenschappers die een sterke negatieve invloed aan de televisie toekenden, sterker geneigd de restrictieve begeleiding toe te passen. Tevens begeleidden deze ouders hun kinderen vaker evaluatief, maar dat gold ook en zelfs sterker voor de ouders die een positieve invloed aan televisie toekenden. Er was geen samenhang tussen de opvatting die de ouders huldigden over de effecten van televisie en de mate waarin zij aan niet-gerichte begeleiding deden.

Recent Nederlands onderzoek met een bescheidener instrumentarium (Van Lil & Vooijs, 1989) wijst op soortgelijke verbanden. Evenals in het Amerikaanse onderzoek zijn ouders die sterke negatieve invloeden aan televisie toeschrijven meer geneigd tot restrictieve begeleidingsactiviteiten, terwijl ouders die de televisie een positieve invloed toedichten vaker evaluatieve begeleidingsactiviteiten toepassen. De gevonden verbanden zijn echter zwakker dan in het onderzoek van Bybee et al. (1982), hetgeen een gevolg kan zijn van het geringe aantal items waarmee tv-begeleiding en opvattingen over de effecten van televisie gemeten zijn. Van Lil en Vooijs (1989) doen dan ook de suggestie het Amerikaanse onderzoek bij een representatieve steekproef van Nederlandse ouders te repliceren.

In het huidige onderzoek herhalen we de studie van Bybee et al. (1982) met als doel na te gaan of de door Bybee et al. (1982) bij massacommunicatie-wetenschappers gevonden vormen van tv-begeleiding (t.w.: restrictieve, evaluatieve en niet-gerichte) ook bij 'gewone' ouders voorkomen. Een tweede doel is na te gaan in welke mate Nederlandse ouders deze vormen van tv-begeleiding toepassen en vast te stellen in hoeverre dit samenhangt met gezinskenmerken en met de effecten die ouders aan het televisie kijken toeschrijven.

Daarbij hebben we het onderzoek van Bybee et al. (1982) enigzins uitgebreid. Aan de vragen over de opvattingen van ouders over tv-effecten op gedrag en attitudes hebben we er enkele over emotionele effecten, zoals angstreacties of onrustig slapen (Surgeon General's Scientific Advisory Committee, 1972; Van Lil, 1988), toegevoegd. Ten aanzien van de vragen over tv-begeleiding hebben we enkele vragen over de niet-gerichte begeleiding toegevoegd, omdat het niet duidelijk is of bij het samen kijken door ouders en kinderen wel pedagogische motieven een rol spelen. Mogelijk komt het samen kijken, met name bij oudere kinderen, vooral door een gemeenschappelijke interesse voor programma's zonder dat van een pedagogisch oogmerk sprake is (Van Lil & Vooijs, 1989).

1.2 *Verwachtingen*

Op basis van de onderzoeken van Bybee et al. (1982) en van Van Lil en Vooijs (1989) verwachten we ten aanzien van de opvattingen van ouders over de effecten van televisie dat:

- de restrictieve begeleiding het meest voorkomt bij ouders die bezorgd zijn over negatieve effecten van televisie;
- de evaluatieve begeleiding eerder toegepast zal worden door ouders met positieve opvattingen over televisie.

De mate waarin ouders tv-opvoedingsactiviteiten toepassen hoeft echter niet alleen af te hangen van de opvattingen die ouders huldigen over de effecten van televisie, maar kan ook afhangen van gezinskenmerken. Wanneer ouders zelf veel tijd aan televisie kijken besteden, is het bijvoorbeeld aannemelijk dat de kans op samen met de kinderen televisie kijken (niet-gerichte begeleiding) toeneemt en dat de kans op restrictieve begeleiding afneemt, terwijl als beide ouders buitenshuis werken de kans op alle vormen van tv-

begeleiding afneemt. Daarnaast zijn blijkens onderzoek (Bower, 1973; Mohr, 1979; Gross & Walsh, 1980; Bybee et al., 1982; Van Lil & Vooijs, 1989) nog diverse andere demografische kenmerken, zoals het geslacht van de ouders, het geslacht en de leeftijd van het kind, het aantal televisie-toestellen dat een gezin bezit en het sociaal economisch milieu, van invloed op het tv-opvoedingsgedrag.

Op grond hiervan is te verwachten dat de restrictieve begeleiding het meest zal voorkomen bij jongere kinderen, bij moeders, bij dochters, bij ouders met een hogere opleiding, wanneer de ouders zelf minder tijd aan televisie kijken besteden en wanneer er minder televisie-toestellen in huis zijn.

De niet-gerichte begeleiding zal het meest voorkomen bij oudere kinderen, bij moeders, bij ouders met een laag opleidingsniveau, wanneer er minder televisie-toestellen in huis zijn, wanneer de ouders zelf meer tijd aan televisie kijken besteden en wanneer er minder thuiswonende kinderen zijn.

Ten aanzien van de evaluatieve begeleiding doen we geen voorspellingen over de samenhang met demografische kenmerken, omdat er soms geen verbanden zijn aangetoond (Bybee et al., 1982), of omdat deze begeleidingsvorm in eerder onderzoek met een te gering aantal items is gemeten.

2 *Methode*

2.1 *Steekproef*

In januari 1989 is door de Afdeling Kijk- en Luister Onderzoek van de NOS bij een representatieve steekproef van ouders (N = 771, gemiddeld 38 jaar) met één of meer thuiswonende kinderen in de leeftijd van 3 tot 18 jaar (gemiddeld 7½ jaar) een telefonisch interview afgenomen.

2.2 *TV-opvoedingsgedrag*

In het interview zijn aan de ouders 16 vragen gesteld over de frequentie ('vaak', 'soms', 'hoogst zelden', 'nooit') waarmee zij tv-opvoedingsactiviteiten toepassen (zie Tabel 1). Specifiek is nagevraagd hoe ouders hun *jongste kind* begeleiden, om vergelijkbare antwoorden over het tv-opvoedingsgedrag van de ouders te verkrijgen. Aan de vragen over tv-opvoeding van Bybee et al. (1982) zijn er twee toegevoegd over samen kijken ('omdat de ouder dat belangrijk vindt' en 'omdat het kind dat wil'). Alle vragen zijn via een gecomputeriseerde rotatie-techniek aan elke ouder in willekeurige volgorde voorgelegd.

2.3 *Gepercipieerde effecten van televisie*

De opvatting van ouders over de effecten die televisie kan hebben op kinderen zijn gemeten door hun uitspraken voor te leggen over diverse invloeden van de televisie die men in de literatuur aantreft. In totaal zijn 20 vragen gesteld; 17 van Bybee et al. (1982) overgenomen uitspraken en drie toegevoegde uitspraken over emotionele effecten (zie Tabel 2). Ook deze uitspraken zijn steeds in willekeurige volgorde aangeboden. De ouders gaven op een vijfpuntsschaal lopend van 'geheel mee eens' tot 'geheel mee oneens' aan in hoeverre zij de uitspraken onderschreven.

2.4 *Demografische kenmerken*

Aan het eind van het interview is gevraagd naar leeftijd en geslacht van de ouders en het jongste kind, gezinsgrootte, opleidingsniveau van de ouders, aantal tv-toestellen in huis en de tijd die ouders aan het kijken naar televisie of video besteden.

3 *Resultaten*

3.1 *TV-opvoedingsgedrag*

De factoriële structuur van een factor-analyse uitgevoerd op de vragen over de frequentie waarmee ouders aan tv-begeleiding doen (zie Tabel 1) vertoont een opmerkelijk grote overeenkomst met de structuur die Bybee et al. (1982) gevonden hebben. Evenals in het Amerikaanse onderzoek onder communicatiewetenschappers vinden we drie factoren die betrekking hebben op respectievelijk de *restrictieve*, *evaluatieve* en *niet-gerichte begeleiding*. Alleen item 10 laadt niet op de factor 'niet-gerichte begeleiding', zoals bij Bybee et al. (1982), maar op de factor 'evaluatieve begeleiding'. Conform onze verwachting laden de twee toegevoegde items over samen kijken het hoogst op de derde factor, welke betrekking heeft op de niet-gerichte begeleiding. De drie factoren met eigenwaarden groter dan 1 verklaren gezamenlijk 44% van de variantie.

De frequenties waarmee ouders 'vaak' of 'soms' tv-opvoedingsactiviteiten toepassen zijn in Figuur 1 weergegeven. Wanneer we alle

Tabel 1 *Factor-analyse; principale componenten-analyse met betrekking tot tv-opvoedingsmethoden; factorladingen na varimax-rotatie*

TV-begeleidingsactiviteiten	I	II	III
Hoe vaak...			
RESTRICTIEVE BEGELEIDING			
1 verbiedt u uw kind naar een bepaald programma te kijken?	.73	.19	-.01
2 beperkt u het televisie kijken van uw jongste kind?	.69	.07	.09
3 zegt u op welke tijdstippen uw kind mag kijken?	.68	-.01	.19
4 geeft u tevoren aan welk programma uw kind mag zien?	.68	.06	.27
5 schakelt u naar een ander net bij ongeschikte programma's	.60	.28	.01
EVALUATIEVE BEGELEIDING			
6 wijst u op bepaalde slechte handelingen in een film?	.10	.68	.12
7 wijst u op bepaalde goede handelingen in een film?	-.00	.65	.30
8 praat u over waarom iemand in een film iets doet?	.08	.65	.26
9 legt u uit dat tv-series en personen verzonnen zijn?	.21	.65	-.07
10 praat u tijdens het kijken over het programma?*	-.02	.46	.38
11 geeft u uitleg over STER-spotjes?	.18	.42	.12
NIET-GERICHTE BEGELEIDING			
12 praat u met uw jongste kind over een programma... ...dat u net bekeken heeft of straks gaat zien?	.10	.26	.64
13 kijkt u samen met uw kind?	-.02	.06	.62
14 kijkt u bewust mee omdat u dat belangrijk vindt?**	.23	.21	.59
15 kijkt u mee omdat uw kind daarom vraagt?**	.11	.18	.58
16 moedigt u aan naar bepaalde programma's te kijken?	.14	.05	.56
Percentage verklaarde variantie	15,9	14,9	13,1

* niet-gerichte begeleiding volgens Bybee et al. (1982)
** niet door Bybee et al. (1982) gestelde vragen

items per factor samennemen, doet 30% van de ouders vaak aan restrictieve begeleiding (33% soms), 27% vaak aan evaluatieve begeleiding (40% soms) en 24% vaak aan niet-gerichte begeleiding (48% soms). Volgens de h-toets van verschillen tussen proporties (Cohen, 1977) is alleen het verschil tussen de mate waarin ouders vaak restrictieve en vaak niet-gerichte activiteiten toepassen significant (h = .135; p = 0,01). De verschillen tussen de mate waarin ouders de evaluatieve begeleiding en de restrictieve of de niet-gerichte begeleiding vaak toepassen zijn niet significant (h is respectievelijk .066 en .069)

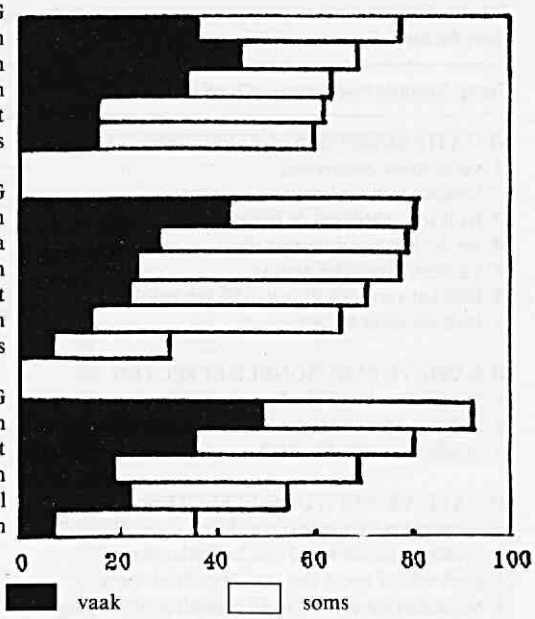
De mate waarin Nederlandse ouders tv-opvoeding toepassen hebben we ook vergeleken met de frequenties waarmee de door Bybee et al. (1982) ondervraagde wetenschappers aan tv-opvoeding doen. Nederlandse ouders passen twee restrictieve, twee evaluatieve en één niet-gerichte begeleidingsactiviteit (items nr. 3, 4, 8, 9 en 13) vaker toe dan de Amerikaanse wetenschappers (h varieert van .173 tot .376; p = 0,05). Drie niet-gerichte activiteiten, één restrictieve en één evaluatieve begeleidingsactiviteit (items nr. 5, 10, 11, 12 en 16) komen yaker bij de Amerikaanse wetenschappers voor (h varieert van .174 tot .686; p = 0,01).

De mate waarin ouders uit de huidige steekproef doen aan tv-opvoedingsactiviteiten laat zich gedeeltelijk ook vergelijken met andere Amerikaanse studies verricht onder representatieve steekproeven van ouders. Bower (1973) stelde exact dezelfde vragen over restrictieve begeleidingsactiviteiten aan 'gewone' ouders (N = 619) met kinderen van 4 tot 12 jaar. Een deelgroep van ouders uit onze studie met vergelijkbare kenmerken (N = 579) verschilt op enkele punten met deze Amerikaanse ouders. Nederlandse ouders geven iets vaker tevoren aan welk programma hun kind mag zien (36% vs. 30% 'vaak'; h = .128, p = 0,05). Amerikaanse ouders spreken daar-

RESTRICTIEVE BEGELEIDING
het televisie kijken beperken
tijdstippen noemen waarop gekeken mag worden
tevoren aangeven wat het kind mag zien
overschakelen naar een ander net
verbieden van bepaalde programma's

EVALUATIEVE BEGELEIDING
uitleggen dat programma's verzonnen zijn
tijdens het kijken praten over een programma
op slechte handelingen in een film wijzen
praten over waarom iemand in een film iets doet
op goede handelingen in een film wijzen
uitleggen geven over STER-spotjes

NIET-GERICHTE BEGELEIDING
samen met het jongste kind kijken
samen kijken omdat de ouder dat belangrijk vindt
praten over een programma voor of na het kijken
samen kijken omdat het jongste kind dat wil
aanmoedigen bepaalde programma's te zien



Figuur 1 *TV-opvoedingsactiviteiten; percentages respondenten die activiteiten vaak en soms toepassen*

entegen vaker een verbod uit over programma's (18% vs. 30% 'vaak'; h = .283, p = 0.01) en schakelen vaker over naar een ander net bij ongeschikte uitzendingen (17% vs. 35% 'vaak'; h = .416, p = 0,01). Op het punt van de evaluatieve begeleiding kan het onderzoek vergeleken worden met een studie van Messaris en Kerr (1983) die aan moeders met kinderen van 6 tot 12 jaar (N = 336) vragen stelden. Nederlandse moeders met even oude kinderen (N = 239) geven vaker uitleg over de onechtheid van televisie (50% vs. 17% 'vaak'; h = .721, p = 0,01), maar praten iets minder vaak over de motieven van tv-personen dan de moeders in het Amerikaanse onderzoek (25% vs. 32% 'vaak'; h = .159, p = 0,05).

3.2 *Gepercipieerde tv-effecten*

Evenals Bybee et al. (1982) vinden we in een factor-analyse uitgevoerd op de uitspraken over de effecten van televisie, een aparte factor voor de positieve invloed van het medium. Blijkens de analyse hoort ook de uitspraak over een verhoogde interesse voor seksualiteit daartoe (zie Tabel 2). In tegenstelling tot Bybee et al. (1982) vormen de uitspraken over gepercipieerde negatieve invloeden van televisie op kinderen niet één, maar drie aparte factoren, die betrekking hebben op: *gedrags-*, *emotionele* en *attitude-effecten*. De vier factoren met eigenwaarden groter dan 1 verklaren gezamenlijk 41% van de variantie.

Gemiddeld over de vijf positieve uitspraken is 58% van de ouders het ermee eens dat de televisie een positieve invloed kan hebben op kinderen. Tegelijkertijd onderkennen ouders ook negatieve invloeden aan het televisie kijken, vooral negatieve emotionele effecten; 79% van de ouders meent dat dat voorkomt. Negatieve invloeden op gedrag en op attitudes worden minder frequent aan televisie toegeschreven (respectievelijk 41% en 16%).

De hier gevonden opvattingen van ouders over de invloeden van televisie op kinderen kunnen we helaas niet met het onderzoek van Bybee et al. (1982) vergelijken. Zij maken geen melding van de mate waarin de communicatie-wetenschappers positieve en negatieve effecten aan televisie toeschrijven.

3.3 *Determinanten van tv-opvoeding*

Via een niet-hiërarchische regressie-analyse hebben we de relatie van het tv-opvoedingsgedrag van ouders met de gepercipieerde effecten van televisie en de demografische kenmerken bepaald (zie Tabel 3). Op basis van de hierboven beschreven factor-analyses zijn drie schalen voor tv-begeleiding en vier schalen voor gepercipieerde televisie-effecten ge-

Tabel 2 *Factor-analyse; principale componenten-analyse met betrekking tot gepercipieerde effecten van televisie; factorladingen na varimax-rotatie*

Gepercipieerde tv-effecten	I	II	III	IV
NEGATIEVE GEDRAGS-EFFECTEN				
1 vermindert creativiteit	.69	-.01	.07	-.10
2 leidt tot minder lezen	.65	.07	-.24	-.05
3 leidt tot onvoldoende lichaamsbeweging	.62	.04	.20	.04
4 verslechtert de concentratie	.54	.19	.17	-.04
5 versterkt agressief gedrag	.44	.37	.19	.04
6 leidt tot vervreemding van de werkelijkheid	.37	.32	.18	-.15
7 leidt tot meer kopen	.35	.10	.19	.29
NEGATIEVE EMOTIONELE EFFECTEN				
8 leidt tot nachtmerries*	.02	.83	.10	.01
9 leidt tot onrustig slapen*	.09	.80	.06	.03
10 maakt kinderen angstig*	.20	.76	.03	.16
NEGATIEVE ATTITUDE-EFFECTEN				
11 versterkt stereotiepe opvattingen over sekserollen	.05	-.02	.58	.07
12 geeft een scheef beeld van buitenlanders	.10	.09	.58	-.05
13 geeft scheef beeld van hoe de politiek werkt	.04	.17	.55	-.05
14 bevordert wens tot directe behoeftebevrediging	.41	-.01	.45	.15
15 geeft een vals idee over wat wel en niet mag	.33	.14	.38	.00
POSITIEVE EFFECTEN				
16 maakt kinderen nieuwsgieriger	-.13	.00	-.06	.65
17 leert kinderen nieuwe woorden	-.07	.13	-.09	.55
18 vergroot de interesse voor seksualiteit**	.25	-.01	.27	.50
19 vergroot de kennis over wat er in de wereld gebeurt	.15	.07	-.40	.44
20 laat kinderen aardiger met elkaar omgaan	-.26	-.18	.23	.37
Percentage verklaarde variantie	12,7	11,5	9,2	7,4

* niet door Bybee et al. (1982) gebruikte uitspraken
** anti-sociale invloed volgens Bybee et al.

construeerd. Elke schaal bestaat uit alle vragen die het hoogst laden op een factor, gewogen volgens de factor-score op die schaal.

Hoewel de verbanden tussen het tv-opvoedingsgedrag van de ouders en de gepercipiëerde effecten van televisie en de demografische kenmerken in de regressie-analyse niet sterk zijn, is er wel sprake van een duidelijk patroon. De restrictieve begeleiding wordt significant vaker toegepast door moeders, bij jongere kinderen, in gezinnen met minder televisie-toestellen in huis en door ouders met een hogere opleiding. Daarnaast zijn ouders die in sterke mate negatieve effecten op het gedrag en de emoties van kinderen waarnemen eerder geneigd tot de restrictieve begeleiding.

De evaluatieve begeleiding laat zich door twee demografische kenmerken, het geslacht van de ouders en hun opleidingsniveau, verklaren. Moeders en ouders met een hoger opleidingsniveau doen meer aan evaluatieve begeleiding. Verder begeleiden ouders die sterk van mening zijn dat de televisie een positieve invloed heeft op kinderen en ouders die vinden dat de emoties van kinderen negatief beïnvloed worden hun kinderen significant vaker evaluatief.

De niet-gerichte begeleiding wordt vaker toegepast door moeders en door ouders die meer opleiding hebben genoten, en komt vaker voor in gezinnen met minder kinderen en in gezinnen waar de ouders zelf meer televisie kijken. Daarnaast begeleiden ouders hun jongste kind vaker niet-gericht wanneer zij er sterker van overtuigd zijn dat de televisie een positieve invloed heeft op de socialisatie van kinderen.

Tabel 3 *Meervoudige regressie-voorspelling van tv-opvoeding door demografische kenmerken en gepercipieerde effecten van televisie*

	Dimensies van TV-begeleiding					
	RESTRICTIEF		EVALUATIEF		NIET-GERICHT	
	r	bèta	r	bèta	r	bèta
Demografische kenmerken						
Leeftijd ouder	-.16***	= = =	-.03	= = =	-.16***	-.04
Geslacht ouder (M = 1, V = 2)	.19***	.18***	.11**	.13***	.15***	.14***
Leeftijd kind	-.25***	-.19***	-.02	.01	-.13**	-.08
Gezinsgrootte	.05	.03	.01	.02	-.13***	-.13***
TV-toestellen thuis	-.18***	-.10**	-.00	= = =	-.07*	-.03
Kijktijd ouder	-.03	= = =	.04	.04	.13***	.09*
Opleiding	.10**	.10**	.06	.11*	.01	.08*
Geslacht kind (M = 1, V = 2)	-.04	-.02	.06	.06	.02	.03
Gepercipieerde effecten						
Positieve effecten	-.04	-.04	.19***	.19***	.14***	.13***
Negatieve emotionele effecten	.20***	.10**	.13**	.09*	.09**	.04
Negatieve gedrags-effecten	.21***	.14***	.05	= = =	.04	= = =
Negatieve attitude-effecten	.12**	.03	.06*	.04	.05	.03
Mult. R		.41***		.27***		.31***

*, ** en *** duiden op significante verbanden op respectievelijk 5%, 1% en 0,1%-niveau; (df = 12,758)

4 *Samenvatting en discussie*

De huidige studie maakt duidelijk dat het tv-opvoedingsgedrag van 'gewone' ouders op dezelfde wijze gestructureerd is als Bybee et al. (1982) bij een selecte groep van Amerikaanse communicatiewetenschappers vonden. Evenals in het Amerikaanse onderzoek maken 'gewone' Nederlandse ouders tegelijkertijd gebruik van de restrictieve, de evaluatieve en de niet-gerichte begeleiding. Slechts één item, dat betrekking heeft op het praten over een programma tijdens het kijken, laadde het hoogst op de factor 'evaluatieve begeleiding', en niet op de factor 'niet-gerichte begeleiding' zoals bij Bybee et al. (1982). Wellicht heeft deze activiteit in de ogen van Nederlandse ouders eerder een evaluatief dan een toevallig, ongericht karakter.

Als het gaat om de frequentie waarmee ouders tv-opvoeding toepassen, lijken 'gewone' Nederlandse ouders iets restrictiever dan Amerikaanse communicatie-wetenschappers, maar minder restrictief dan 'gewone' Amerikaanse ouders. Deze verschillen zijn echter moeilijk te verklaren, omdat diverse factoren een rol spelen. Bybee et al. (1982) hebben een bijzondere groep van hoog opgeleide wetenschappers onderzocht en bovendien bestond hun steekproef voor het merendeel uit mannen (86%), welke in de regel minder aan tv-opvoedingsactiviteiten doen dan vrouwen. Ook de vergelijking met de studie van Bower (1973) bij 'gewone' Amerikaanse ouders moeten we met reserve bezien, omdat dat onderzoek reeds meer dan 16 jaar geleden is uitgevoerd. Het lijkt daarom niet zinvol conclusies te binden aan de verschillen in het opvoedingsgedrag voor zover het de restrictieve begeleiding betreft.

Wel interessant zijn de uitkomsten dat de niet-gerichte begeleiding vaker voorkomt bij de Amerikaanse wetenschappers, terwijl de evaluatieve activiteiten iets vaker door 'gewone' Nederlandse ouders dan door Amerikaanse wetenschappers worden toegepast. Bovendien komt het uitleg geven over de onechtheid van televisie-programma's (evaluatief) vaker voor bij Nederlandse moeders dan bij Amerikaanse. We kunnen aannemen dat het hier om een cultureel verschil tussen de Verenigde Staten en Nederland gaat. Wellicht is het voor Amerikaanse ouders relatief moeilijker de evaluatieve begeleiding toe te passen, omdat

er in de Verenigde Staten een groter aanbod van programma's is en de televisie in het algemeen vaker aan staat. Misschien wordt er in Amerikaanse gezinnen meer over televisie in het algemeen gepraat dan in Nederlandse gezinnen, terwijl Amerikaanse ouders tijdens het kijken toch minder uitleg geven aan hun kinderen over de inhoud. Een replicatie van het huidige onderzoek bij een representatieve steekproef van Amerikaanse ouders is echter gewenst om het onderscheid tussen de drie vormen van het tv-opvoedingsgedrag van Nederlandse en Amerikaanse ouders duidelijk te maken.

Evenals in het onderzoek van Bybee et al. (1982) en van Van Lil en Vooijs (1989) denken de ondervraagde Nederlandse ouders over de rol die televisie in het socialisatieproces van kinderen speelt meerdimensionaal. De meeste ouders menen dat aan het televisie kijken van kinderen zowel voor- als nadelen zijn verbonden. Opvallend is evenwel dat de door televisie verhoogde interesse van kinderen voor seksualiteit in onze studie als een positieve invloed wordt opgevat. In het onderzoek van Bybee et al. (1982) zagen de wetenschappers de verhoogde interesse voor seksualiteit als een negatieve invloed van televisie. Wellicht zijn de verschillende opvattingen wederom te verklaren als culturele verschillen. In het algemeen is de invloed van een 'Moral Majority' in de Verenigde Staten veel sterker dan in Nederland. Het beleid van Nederlandse omroeporganisaties ten aanzien van programma's over liefde en erotiek is dan ook vrijzinniger, en de houding van Nederlanders ten aanzien van seksualiteit en zeden is in de loop der jaren eveneens sterk geliberaliseerd (Sociaal Cultureel Planbureau, 1988).

De door ons toegevoegde uitspraken over emotionele effecten van televisie blijken van groot belang te zijn. De drie uitspraken vormen een nieuwe zelfstandige factor. Voorts vindt een grote meerderheid van de ouders dat de televisie een negatieve invloed heeft op de emoties van kinderen. Ouders zijn daar vaker van overtuigd dan van de andere effecten. Bovendien blijken ouders die bezorgd zijn over negatieve emotionele effecten van televisie, niet alleen vaker de restrictieve begeleiding toe te passen, maar ook de evaluatieve begeleiding. Kennelijk willen ouders de angstreacties als gevolg van het televisie kijken bij kinderen niet alleen via restricties op het kijkgedrag tegengaan, maar hechten ze er ook belang aan tijdens het kijken programma-inhouden aan hun kinderen uit te leggen.

Eveneens nieuw is dat bij ouders die overtuigd zijn van het 'goede' van televisie, niet alleen vaker de evaluatieve begeleiding voorkomt, zoals we verwachtten, maar dat deze ouders hun kinderen ook vaker niet-gericht begeleiden. Wellicht vinden we dit verband met de niet-gerichte begeleiding, omdat we specifieke vragen over de beweegredenen van ouders om samen met de kinderen televisie te kijken toegevoegd hebben.

In het huidige onderzoek hadden we geen informatie ter beschikking over het al of niet buitenshuis werken van de ouders, het aangesloten zijn op het kabelnet en het kijkgedrag van het kind. Desalniettemin geven de wel onderzochte kenmerken een interessant inzicht in de mate waarin tv-opvoeding bij Nederlandse ouders voorkomt. Van de demografische kenmerken zijn het opleidingsniveau en het geslacht van de ouders de meest belangrijke voorspellers voor het tv-opvoedingsgedrag. Moeders en ouders uit hogere sociale milieus passen niet alleen het meest de restrictieve en de niet-gerichte begeleiding toe, maar begeleiden hun kinderen ook vaker evaluatief. Kennelijk vinden ouders met een hogere opleiding en moeders, die traditioneel meer dan vaders bij de opvoeding van kinderen betrokken zijn, het belangrijker kinderen bij het televisie kijken te begeleiden en zetten dat dan ook om in feitelijk gedrag.

Ten aanzien van de voorspellingen over de samenhang van het tv-opvoedingsgedrag van ouders met demografische kenmerken worden er slechts enkele niet bevestigd. In tegenstelling tot de bevindingen van Gross en Walsh (1980), passen Nederlandse ouders de restrictieve begeleiding niet vaker toe bij dochters dan bij zonen. Evenmin zijn ouders die zelf meer televisie kijken minder restrictief dan ouders die minder tijd aan televisie kijken besteden. Ouders zijn wel minder restrictief naarmate de kinderen ouder worden en er meer televisie-toestellen in huis zijn. Kennelijk vinden ouders het in die situaties niet langer noodzakelijk hun kinderen te beschermen tegen negatieve invloeden van televisie.

In tegenstelling tot onderzoek van Gross en Walsh (1980) en van Bybee et al. (1982) hangt de niet-gerichte begeleiding noch met het aantal televisie-toestellen in het gezin, noch met

de leeftijd van de kinderen samen. De niet-gerichte begeleiding is meer van 'toevallige' situationele kenmerken afhankelijk en is waarschijnlijk dan ook de minst gemotiveerde vorm van tv-begeleiding. Het is voor ouders geen bijzonder grote opgave met de kinderen samen televisie te kijken, wanneer zij zelf ook veel kijken of wanneer er maar weinig kinderen in het gezin zijn.

De evaluatieve begeleiding wordt opmerkelijkerwijs nauwelijks voorspeld door demografische kenmerken, maar hangt vooral met de opvattingen van ouders over effecten van televisie samen. Het huidige onderzoek ondersteunt daarmee de conclusie van Bybee et al. (1982) dat de evaluatieve begeleiding de meest geïnvolveerde vorm van tv-begeleiding is. Wanneer ouders hun kinderen evaluatief begeleiden en daarbij zowel oog voor de positieve, alsook voor de negatieve kant van televisie hebben, is de kans het grootst dat kinderen kritisch worden over programma-inhouden en dat zij deze op waarde weten te schatten.

De wijze waarop ouders zich via verschillende tv-begeleidingsactiviteiten bemoeien met het kijkgedrag van hun kinderen kan de positieve (leer-)effecten van televisie versterken, alsook de negatieve (angst- en gedrags-) effecten tegengaan. Praktisch gezien wijzen de resultaten van het huidige onderzoek er dan ook op dat wanneer we ouders voor de evaluatieve of voor de niet-gerichte begeleiding willen interesseren, we hen het best kunnen overtuigen van de positieve effecten van televisie. Ouders op de negatieve effecten op gedrag en emoties van kinderen wijzen, lijkt daarentegen een betere strategie hen tot de restrictieve begeleiding aan te zetten.

De mate waarin deze strategieën effectief kunnen zijn zouden via de veel georganiseerde ouderavonden op basisscholen over 'kind en televisie' systematisch onderzocht kunnen worden, door de ouders over de negatieve of over de positieve kanten van televisie te informeren en hen voor en na zo'n avond te ondervragen over de mate waarin zij de verschillende vormen van tv-opvoeding toepassen. Tevens suggereren de uitkomsten van het huidige onderzoek dat bij de initiatieven om ouders aan te sporen tot tv-opvoeding de meeste nadruk op moeders gericht zou kunnen worden, omdat zij daar toch al het meest toe geneigd zijn. Wellicht kan bekeken worden of en hoe zij via artikelen in opvoedings- en damesbladen, tv-gidsen, of via op hen gerichte televisie-programma's gestimuleerd kunnen worden tot tv-opvoeding.

Literatuur

Ball, S. & G.A. Bogatz, *The first year of Sesame Street: an evaluation*. Princeton, N. J.: Educational Testing Service, 1970.

Barcus, F. E., Parental influence on children's television viewing. *Television Quarterly*, 1969, *8*, 63-73.

Bower, R. T., *Television and the public*. New York: Holt, Rinehart and Winston, 1973.

Bybee, C., D. Robinson & J. Turow, Determinants of parental guidance of children's television viewing for a special subgroup: Mass media scholars. *Journal of Broadcasting*, 1982, *26*, 697-710.

Cantor, J. & B. J. Wilson, Modifying fear responses to massmedia in preschool and elementary school childeren. *Journal of Broadcasting*, 1984, *28*, 431-443.

Cohen, J., *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1977.

Comstock, G.A., S. Chaffee, N. Katzman, M. McCombs & D. Roberts, *Television and human behavior*. New York: Columbia University Press, 1978.

Corder-Bolz, C. R., Mediation: the role of significant others. *Journal of Communication*, 1980, *30*, 106-118.

Corder-Bolz, C. R., Television literacy and critical television viewing skills. In: D. Pearl, L. Bouthilot & J. Lazar (Eds.). *Television and behavior: Ten years of scientific progress and implications for the eighties* (pp. 91-101). Washington, D.C.: U.S. Government Printing Office, 1982.

Gadberry, S., Television as a babysitter: A field comparison of preschoolers' behavior during playtime and during television viewing. *Child Development*, 1974, *45*, 1132-1136.

Gross, L. S. & R. P. Walsh, Factors affecting parental control over childeren's television viewing: A pilot study. *Journal of Broadcasting*, 1980, *24*, 411-419.

Lil, J. E. van, *Rapport inzake een onderzoek naar de betekenis van de massamedia voor leerlingen van de lagere school*. Hilversum: NOS, Afdeling Kijk- en Luisteronderzoek, 1978.

Lil, J. E. van, *Beeld van een kind*. Hilversum: Nos, Afdeling Kijk- en Luisteronderzoek, 1988.

Lil, J. E. van & M. W. Vooijs, TV-opvoeding in het gezin. *Massacommunicatie*, 1989, *17*, 36-47.

McLeod, J.M., M.A. Fitzpatrick, C.J. Glynn & S.F. Fallis, Television and social relations: family influences and consequences for interpersonal behavior. In: D. Pearl, L. Bouthilot & J. Lazar (Eds.), *Television and behavior: Ten years of scientific progress and implications for the eighties* (pp. 272-287). Washington, D.C.: U.S. Government Printing Office, 1982.

Messaris, P., Family conversations about television. *Journal of Family Issues*, 1983, *4*, 293-308.

Messaris, P. & D. Kerr, Mothers comments about tv: Relation to family communication patterns. *Communication Research*, 1983, *10*, 175-194.

Meyer, G. de, A. Hendriks & G. Fauconnier, Een ontwenningskuur tegen televisieverslaving. *Communicatie*, 1985, *15*, 1-15.

Mohr, P.J., Parental guidance of children's viewing of evening television programs. *Journal of Broadcasting*, 1979, *23*, 213-228.

Nikken, P., *Mag de tv aan? Kinderen en televisie*. Amersfoort: ACCO/de Kinderkast, 1989.

Palmer, E.L., A.B. Hockett & W.W. Dean, The television family and children's fright reactions. *Journal of Family Issues*, 1983, *4*, 279-292.

Sociaal en Cultureel Planbureau. *Sociaal en Cultureel Rapport 1988*. Alphen aan de Rijn: Samsom, 1988.

Surgeon General's Scientific Advisory Committee on Television and Social Behavior. *Television and growing up. The impact of television violence*. Washington, D.C.: Government Printing Office, 1972.

Vooijs, M.W., *TV-opvoeding*. Lisse: Swets & Zetlinger, 1987.

Voort, T.H.A. van der (Ed.), Media-opvoeding: Theorie, prakijk en onderzoek. *Massacommunicatie*, 1981, *17*, 6-117.

Curricula vitae

P. Nikken (1960) studeerde psychologie (1986) en is als coördinator van het Kinderkast-project werkzaam voor de Stichting Kinderpostzegels Nederland.

Adres: Stichting Kinderpostzegels Nederland, de Kinderkast, Stationsweg 25, 2312 AS Leiden.

J. E. van Lil (1941) studeerde sociologie (1970) en is als media-onderzoeker werkzaam bij de Afdeling Kijk- en Luisteronderzoek van de NOS.

Adres: NOS, Afdeling Kijk- en Luisteronderzoek, Postbus 444, 1200 JJ Hilversum.

Manuscript aanvaard 4-4-'90

Summary

Nikken, P. & J.E. van Lil. 'Determinants of parental quidance of children's tv viewing: a replication study.' *Pedagogische Studiën*, 1990, *67*, 417-426.

By means of a telephonic interview a representative sample of 771 parents with children from three to eighteen years answered questions concerning the views parents hold about the effects of television on children, and about how often parents guide their children's television viewing. Conform an American study we replicated (Bybee, Robinson & Turow, 1982), parental guidance can be divided in three distinct forms: restrictive, evaluative, and unfocused guidance. Parents are more restrictive when they hold negative views about television effects, with younger children, and when there are less television sets in the house. Parents exert more evaluative guidance when they hold strong positive or negative views about television effects. Parents exert more unfocused guidance when they hold positive views about television effects, in smaller families and when they watch more television themselves. Mothers and higher educated parents exert all forms of guidance more frequently.

Boekbespreking

Elly Singer, *Kinderopvang en de moeder-kindrelatie.* Van Loghum Slaterus bv., Deventer, 1989, 310 pag., *f* 48,50, ISBN 90 368 0165 6.

De aanleiding voor dit proefschrift was, zo schrijft Singer in haar Inleiding, de feministische strijd voor betaalbare kinderopvang die vanaf de jaren zeventig werd gevoerd, en de – verontwaardigde – reacties van opvoedingsdeskundigen hierop. Singer benoemt deze confrontatie in termen van moedergerichte contra kindgerichte opvattingen en vraagt zich af of deze twee opvattingen per definitie tegengesteld zijn. Om een antwoord op deze vraag te kunnen formuleren gaat ze op zoek naar de historische wortels van de hedendaagse situatie.

Het historische overzicht begint omstreeks 1800, wanneer op meerdere plaatsen in Europa en de Verenigde Staten – in het kader van de armenzorg – bewaarscholen werden opgericht. De vroege opvoeding in de bewaarschool werd gezien als een van de mogelijkheden om criminaliteit en zedeloosheid uit te bannen en om daarmee mogelijke bedreigingen voor de stabiliteit van de samenleving weg te nemen. Bovendien hadden de 'welingerichte bewaarscholen' ook een functie als voorbereiding voor het volksonderwijs, dat in deze periode de klassikale vorm kreeg. Daarnaast speelde de opkomst van de moederschaps- en gezinsideologie een rol. Het liefdadigheidswerk in de bewaarscholen werd gezien als een uitbreiding van de moederlijke zorg van vrouwen uit de burgerij.

Begin negentiende eeuw was de bewaarschool, aldus Singer, een sociaal 'lege' ruimte. Er waren geen pedagogische tradities in het werken met groepen jonge kinderen; de theoretische en methodische leegte tussen gezinsopvoeding en onderwijs moest nog worden gevuld. Na een aanzet door Pestalozzi kreeg de opvoeding van jonge kinderen in de tweede helft van de negentiende eeuw een belangrijk impuls door de theorieën van Fröbel. Fröbel was medewerker van Pestalozzi geweest, wat blijkt uit vergelijkbare opvattingen over de ideale moederliefde, waarin moeders van vlees en bloed tekort schoten. Fröbel was de eerste pedagoog die met een uitgewerkte methode kwam, de 'Kindergarten', waar kinderen tussen één en zeven jaar onder leiding van speciaal opgeleide vrouwen konden spelen. Om de vrouwen op te leiden werd een 'Hochschule für Frauen' opgericht. Het gevolg van deze en andere opleidingen, namelijk dat vrouwen deze gebruikten om hun zorgend leven buiten de familiegrenzen uit te breiden, stond Fröbel weinig aan.

Veel meer dan in Europa, sloeg de Kindergartenbeweging aan in de Verenigde Staten. Daar ontstond aan het einde van de negentiende eeuw de eerste 'child movement' met als bekendste psycholoog G. Stanley Hall. Hij verdedigde de opvatting dat kinderopvoeding gebaseerd moet worden op empirische kinderstudies, en niet op metafysische beschouwingen. Het geloof dat door een ideale kinderopvoeding de samenleving verbeterd kon worden, bleef echter behouden. Zijn ideale opvoeding baseerde Hall op de recapitulatietheorie: het kind zou in zijn ontwikkeling de ontwikkeling van de soort herhalen. Daaruit volgde volgens Stanley Hall dat de allerjongste kinderen bij moeder thuis hoorden. In de fase van poëtische fantasie en mythisch denken was de Kindergarten de ideale plaats, en tussen acht en twaalf jaar zouden kinderen behoefte hebben aan discipline en spanning. Voor oudere kinderen was de feminisering van het onderwijs, met zijn slapte en sentimentaliteit, voor Hall een gruwel. Na de eerste Wereldoorlog ontstond een tweede 'child movement'. Binnen onderzoeksinstituten werden 'nursery schools' opgericht, waar kinderen tussen anderhalf en vier jaar geobserveerd konden worden. Kindercentra dienden als middel om de opvoeding door moeders thuis op een hoger plan te brengen. De competentie van moeders werd betwijfeld; hun betrokkenheid werd te emotioneel geacht.

In de periode na de Tweede Wereldoorlog onderscheidt Singer twee belangrijke ontwikkelingen. De Engelse psychiater Bowlby stelde dat 'maternal deprivation' of een verstoorde emotionele band tussen moeder en kind onherstelbare schade voor het kind, én voor de samenleving tot gevolg had. Buitenshuis werken van moeders raadde hij af; psychologische hulp en begeleiding van moeders was zijns inziens erg belangrijk. Moeders werden in hoge mate verantwoordelijk gesteld voor het herstel van een maatschappelijke orde, waarvan zijzelf buitengesloten werden. In Bowlby's latere attachmenttheorie kwam veilige hechting centraal te staan. Volgens hem had het kind een aangeboren neiging om de bescherming van volwassenen te zoeken. Daaraan tegengesteld was een tweede aangeboren neiging van exploratief gedrag in de buitenwereld. Als de moeder goed op het kind reageerde, ontstond een veilige gehechtheid van waaruit de omgeving kon worden verkend. In de jaren tachtig introduceerde in ons land Van IJzendoorn c.s. de zogeheten 'verbredingshypothese', die inhield dat kinderen meerdere gelijkwaardige gehechtheidsrelaties kon-

den vormen.

De tweede naoorlogse ontwikkelingslijn die Singer beschrijft is het Amerikaanse Head-Start project uit het midden van de jaren zestig, bedoeld om de opvoeding van 'achtergestelde' of 'cultureel gedepriveerde' kinderen ter hand te nemen. Hierin werd vooral gezocht naar mogelijke positieve effecten van kinderopvang op de cognitieve ontwikkeling; onderzoek naar mogelijke gevolgen van moeder-kind scheiding bleef – ondanks het bestaan van Bowlby's theorieën – achterwege. Wel hadden de tegenvallende resultaten van de kindercentra die binnen het project fungeerden tot gevolg dat ouder/moedereducatie krachtig werd gestimuleerd.

Samenvattend constateert Singer dat er in de geschiedenis drie belangrijke motieven voor het ontstaan van kinderopvang waren: het opvoeden van probleemkinderen teneinde sociale onrust, vooral in de lagere sociale milieus, te voorkomen; de voorbereiding op het lager onderwijs; en het uitdragen van een nieuw opvoedingsideaal in met name de middle class. In pedagogische en psychologische theorieën over kinderopvang werd (en wordt) over het algemeen de moeder als belangrijkste opvoedster van jonge kinderen gezien. Omdat moeders dat niet goed (lagere sociale klassen) of niet zò goed (middle class) konden, werd een aantal aspecten van de opvoeding overgenomen door deskundigen. Het loskoppelen van vrouw zijn en opvoeden, en de strijd voor een financiële beloning voor opvoeden kwam pas in jaren zeventig. Om de opvoedingsproblematiek opnieuw te doordenken is het volgens Singer noodzakelijk om in te zien dat het sociale isolement van moeder en kind en de sterke emotionele banden tussen beiden de bron is van veel opvoedingsonzekerheid, dat het geïsoleerde gezin machtsmisbruik van ouders tegenover kinderen onzichtbaar houdt, en dat we tegenstelling tussen beschermend gezin en boze buitenwereld moeten doorbreken. Waarom, vraagt Singer, mijns inziens terecht, zouden kinderen niet mogen weten dat vaders en moeders buiten hen om ook andere belangrijke dingen te doen hebben?

De sterkte van Singers proefschrift is tegelijkertijd de zwakte ervan. De auteur neemt ons mee op een boeiende tocht langs mogelijke bronnen van het denken over kinderopvang. Soms blijkt een bron echter niet zo terzake; soms blijkt een tevoren niet aangekondigd uitstapje nodig te zijn. Bovendien zijn er nogal wat herhalingen en overlappingen in de tekst, en worden in de loop van het historische verhaal soms conclusies getrokken of aanbevelingen gedaan over het heden. Door deze aanpak wordt het doorzettingsvermogen van de lezer/es op de proef gesteld. Ondanks deze soms wat onheldere structurering geeft het boek een schat van historische informatie, die niet alleen de geschiedenis van het denken over kinderopvang omvat, maar ook de verwevenheid van het denken over vrouwelijkheid en vrouwen, maatschappelijke beschaving en kinderopvoedig. Het geeft op die manier een welkome verbreding van ontwikkelingslijnen die in andere publikaties op verwante gebieden aangegeven zijn. De conclusie die uit het boek naar voren komt, namelijk dat de geschiedenis niet veel aanknopingspunten geeft om kindgericht en moedergericht denken te combineren, had wat mij betreft wel wat explicieter geformuleerd mogen worden. We staan voor een nieuw pedagogisch experiment, zo toont de dissertatie van Singer aan. Voor het doordenken van dit experiment geeft Singer een aantal zinnige aanknopingspunten.

M. Lunenberg

Kunnen jonge kinderen 'Ordenen'?
Onderzoek naar deelvaardigheden voorbereidend rekenen van 4-6 jarigen

J. J. VAN KUYK
Cito, Instituut voor Toetsontwikkeling, Arnhem

Samenvatting

De vraag wat kinderen kunnen is in het kleuteronderwijs nog nauwelijks aan de orde gesteld.

In het verlengde van een onderzoek naar de kwaliteit van een toetsprogramma met als onderwerp 'Ordenen', een basis-vaardigheid op het gebied van waarnemen, denken, taal, getalbegrip en de voorbereiding op het rekenen, werd de vraag gesteld: wat kunnen kinderen op het gebied van 'Ordenen'. Deze vraag heeft belang voor verschillende aspecten van het kleuteronderwijs. Eén ervan is dat er belangrijke momenten in de ontwikkeling kunnen worden vastgesteld, ook wel 'mijlpalen' genoemd, aan de hand waarvan achterstanden kunnen worden vastgesteld, met het oog op een gerichte hulpverlening.

De vraag werd gespecificeerd in een aantal deelvragen die betrekking hebben op verschillen in achtergrondkenmerken en verschillen tussen hoog- en laagscorenden.

Uit de resultaten blijkt dat het mogelijk is didactisch hanteerbare 'mijlpalen' aan te geven. Tevens blijkt dat de verschillen tussen kinderen met betrekking tot het bereiken van deze 'mijlpalen' groot zijn. Sommige kinderen blijken de meeste 'mijlpalen' aan het eind van het kleuteronderwijs niet bereikt te hebben; ze hebben grote achterstanden, terwijl andere kinderen al bij de intrede in de basisschool een groot aantal 'mijlpalen' bereikt hebben. In de discussie wordt gepleit voor speciale aandacht voor zowel kinderen met achterstanden als voor kinderen die al vroeg de 'mijlpalen' hebben bereikt. Er wordt gepleit voor verbreding van het aanbod en voor een belangrijke plaats van de evaluatiefunctie in het kleuteronderwijs.

1 *Inleiding*

De discussie over de vraag wat het onderwijsaanbod zou moeten zijn voor de eerste twee groepen van de basisschool, het voormalige kleuteronderwijs, is met de komst van de basisschool in 1985 toegenomen. Uit een rapport van de inspectie (O & W, 1989) blijkt, dat de integratie organisatorisch redelijk tot stand gekomen is, maar inhoudelijk nog zeer te wensen overlaat. Niet alleen wat er onderwezen dient te worden, het onderwijsaanbod, ook wat het resultaat van het onderwijs moet zijn, moet in deze discussie worden betrokken. Op het gebied van het schoolrijpheidsonderzoek zijn er in het verleden criteria opgesteld wat kinderen zouden moeten kunnen aan het begin van de lagere school (Dumont & Kok, 1970). Wat er tijdens het kleuteronderwijs geleerd werd, was geen voorwerp van onderzoek. Later toen de zorg voor kinderen met achterstanden sterk op de voorgrond kwam, we denken hierbij aan het 'Head Start' programma in de Verenigde Staten en de opkomst van de compensatieprogramma's in ons land, kwam het curriculum van het kleuteronderwijs meer in het vizier. Ontwikkeling hoefde niet uitsluitend te worden afgewacht, maar kon ook worden bevorderd gedurende het kleuteronderwijs (Groenendaal, 1978; Van Kuyk, 1985; Stallings & Stipek, 1986; Slavin & Madden, 1989). Over de resultaten van de compensatie- en stimuleringsprogramma's bestaat geen eenstemmigheid. In het algemeen zijn de verwachtingen niet hoog gestemd, maar tegenover een niet gerealiseerde hooggestelde verwachting (Slavenburg, 1990) staat de geruststellende mededeling van Dronkers (1990) op de pessimistische visie van Slavenburg dat er tegen de dood ook nog geen medicijn bestaat. Aanvankelijke korte-termijnsuccessen direct na de training bleken op de langere termijn van twee of meer jaren niet houdbaar. In de volwassenheid echter bleken

kleuterschoolprogramma's toch hun uitwerking te hebben gehad, blijkens betere carrièremogelijkheden, grotere zelfstandigheid en minder crimineel gedrag in vergelijking met volwassenen die een dergelijk programma niet volgden (Weikart, 1987). Over het algemeen blijken vooral effecten op korte termijn gemeld te kunnen worden (Slavenburg, 1989). Afgezien van de resultaten is wel gebleken dat niet alle leerlingen met gelijke kansen aan het kleuteronderwijs beginnen. Door factoren in de woonomgeving, in het gezin, in de schoolsituatie en in het kind zelf (ARBO, 1988) zijn er leerlingen in het kleuteronderwijs die belangrijke kenmerken missen die voor het leren op school van belang zijn. We kunnen daarbij denken aan een geringe zelfsturing, passiviteit, een slechte selectieve aandacht, een gebrekkige analyse en structurering, moeite met abstraheren en generaliseren en een snelle vermoeibaarheid (Van der Leij, 1985; Van Parreren, 1989). Op grond hiervan kunnen er tussen kinderen grote verschillen in ontwikkeling ontstaan, die in het basisonderwijs onvoldoende worden opgevangen. Aan leerlingen die onvoldoende geëquipeerd zijn zal speciale zorg besteed moeten worden.

Hoewel de resultaten vooralsnog niet bevredigend zijn, worden vooral goed gestructureerde programma's aanbevolen, waarin metacognitieve componenten vertegenwoordigd zijn (Slavenburg, Leune, Creemers, Peters, 1989). Cunningham (1989) toonde in dit verband aan, dat voorbereidende leesoefeningen in het kleuteronderwijs succesvol kunnen zijn voor het leren lezen, maar dat oefeningen die gepaard gaan met een metacognitieve benadering significant meer resultaat opleveren, dan alleen 'drill and practice'. Van belang is dus dat de leerling zich bewust wordt van het nut van bepaalde oefeningen.

Slavin en Madden (1989) tonen aan dat een tutoring-benadering, waarbij de leerkracht in een één-éénsetting lesgeeft het meest succesvol is. Het programma dient echter omvattend te zijn, een handleiding te hebben, intensief te zijn, directe instructie te bevatten en er dient regelmatig geëvalueerd te worden. Voorzieningen van vroege diagnostiek en behandeling horen eveneens tot belangrijk geachte voorzieningen (Van Kuyk, 1982; Van der Leij, 1985; Stevens, 1987). Hierin ligt de erkenning, dat niet alleen het vroegere lager onderwijs, maar ook het kleuteronderwijs een belangrijke bijdrage levert aan de ontwikkeling van het jonge kind. Is deze bijdrage voldoende als er een bepaald aanbod is, aan de hand waarvan het kind zich kan ontwikkelen of is het ook wenselijk de resultaten van het gegeven onderwijs vast te stellen?

Over de vraag wat inhoudelijk in de eerste twee groepen van de basisschool het aanbod moet zijn, is internationaal niet zoveel verschil van mening, de laatste paar jaar zelfs niet meer als het gaat om het aanleren van de vaardigheden lezen, schrijven en rekenen (Van Kuyk, 1987; ARBO, 1990; The early years curriculum group, 1989). Wat betreft de voorbereiding op deze vaardigheden is er een duidelijke trend deze in het curriculum van de eerste twee groepen van de basisschool op te nemen.

Als we echter de vraag stellen wat het feitelijk resultaat is of zou dienen te zijn van twee jaar kleuteronderwijs en of alle leerlingen een bepaald resultaat behalen, dan is daarop geen passend antwoord mogelijk. Er zijn wel talrijke ontwikkelingspsychologische beschrijvingen wat kinderen op bepaalde leeftijden kunnen, maar onderwijskundig in termen van doelen en resultaten van onderwijs is deze vraag nog nauwelijks aan de orde gesteld.

Resultaten van onderwijs, zo is onze stelling, zouden een basis kunnen vormen voor de verheldering van opvattingen en de didactische wegen die in het kleuteronderwijs bewandeld kunnen worden. Resultaten vatten we hier ruim op. We bedoelen ermee prestaties van leerlingen zonder de vraag te stellen of deze tot stand zijn gekomen door spontane ontwikkelingsprocessen dan wel door gericht gestuurde onderwijsleerprocessen.
We noemen hier enkele punten:

- Resultaten kunnen inzicht geven wat kleuters kunnen en wat van hen verwacht mag worden. Voorkomen kan dan worden dat van kinderen teveel of te weinig gevraagd wordt. Overhaasting, waarvoor met name door ontwikkelingspsychologen gewaarschuwd wordt (Elkind, 1981), kan dan door het beschikbare cijfermateriaal voorkomen worden; daarnaast kan ook onnodige vertraging of 'pas op de plaats' vermeden worden.
- Resultaten kunnen aanknopingspunten bieden voor curriculumontwikkeling. Met name het reguleren van het aanbod en het structureren ervan kan mede op basis van resultaten geschieden. Er is dan een empi-

risch fundament voor onderwijskundig handelen.

- Aan de hand van resultaten kunnen 'mijlpalen' in de ontwikkeling worden vastgesteld, waarvoor de ARBO (1990) sterk pleit. Onder een 'mijlpaal' verstaan we een belangrijk moment in de ontwikkeling dat in termen van evaluatie of diagnose interpreteerbaar is en in termen van te geven onderwijs hanteerbaar. Onafhankelijk van de verschillende opvattingen over kleuteronderwijs is het van belang te weten of een kind zich volgens een bepaald patroon ontwikkelt of het achterblijft bij de verwachtingen of juist sterke vooruitgang boekt. Aan de hand hiervan kan de evaluatiefunctie worden ontwikkeld in het kleuteronderwijs en kan het onderwijsproces adequater gestuurd worden.
- Resultaten kunnen tenslotte bijdragen aan vroegtijdige onderkenning van achterstanden en leermoeilijkheden. Op basis daarvan kan worden nagegaan of het gewenst is maatregelen te nemen voor een gerichte hulpverlening.
  In een wat verder verwijderd perspectief zou kunnen worden gedacht aan kwaliteitscontrole van het kleuteronderwijs, met name wanneer resultaatmetingen herhaald zouden worden. Zonder de discussie over opvattingen te voeren, zijn er voldoende redenen om in het kleuteronderwijs onderzoek te doen naar de resultaten van ontwikkeling en onderwijs.

In het onderhavige onderzoek beperken we ons tot een domein dat we beschouwen als een basisvaardigheid voor het huidig en toekomstig leren in de basisschool en dat in de praktijk van het Nederlands onderwijs een plaats heeft. Het betreft 'Ordenen', een brede algemene vaardigheid, die bijdraagt aan de cognitieve ontwikkeling, met name aan waarnemen, denken en taal en aan de ontwikkeling van het getalbegrip en het leren rekenen, een van de hoofd-doelstellingen van het basisonderwijs. Dit domein is uitgewerkt in een programma met een elftal onderdelen, dat tot stand is gekomen op basis van theoretische inzichten en een analyse van een aantal praktijkprogramma's (Van Kuyk, 1990).

2 *Vraagstelling*

De vraagstelling luidt derhalve: wat presteren kleuters van 4-6 jaar op het gebied van 'Ordenen'? Deze vraag verdelen we in een aantal specifieke vragen:

- Hoe ontwikkelt deze vaardigheid zich gedurende de kleuterperiode? Op grond hiervan zouden 'mijlpalen' in de ontwikkeling van het 'Ordenen' kunnen worden vastgesteld.
- Wat zijn de verschillen tussen groepen leerlingen wat hun achtergrondkenmerken betreft. Hierbij zijn variabelen gekozen waarvan er enkele ook in PPON-onderzoek aan bod komen (Wijnstra, 1988).
- Wat zijn de verschillen tussen leeftijdsgroepen?
- Zijn er verschillen tussen jongens en meisjes?:
  presteren jongens gemiddeld beter op de taken van 'Ordenen' dan meisjes? Jongens blijken betere rekenprestaties te leveren in het midden en einde van de basisschool, zoals uit PPON-onderzoek blijkt (Wijnstra, 1988). Is dat ook het geval bij Ordenen, waarin een aantal aspecten van het voorbereidend rekenen vertegenwoordigd zijn?
- Doet het aantal maanden genoten onderwijs ertoe of maakt het niet uit of kinderen langer of korter naar school gaan? Uit onderzoek in de Verenigde Staten (Stallings & Stipek, 1986) bleek dat kinderen die een hele dag naar school gingen beter presteren dan kinderen die slechts een halve dag naar school gaan, zij het dat de resultaten van korte duur waren. In ons land zou de vraag gewettigd zijn in hoeverre een vierjarigen maatregel schadelijk is voor de leerlingen.
- Ten slotte werd de vraag gesteld of de status van de school uitgedrukt in gemiddelde leerlinggewichten verschillen veroorzaakt, zoals dat ook in het PPON-onderzoek het geval is.
  Scholen met een groot aantal arbeiders- en allochtone kinderen scoren lager dan scholen met relatief weinig van deze kinderen (Wijnstra, 1988).
  In ons onderzoek hebben we onderscheid gemaakt tussen scholen met relatief veel arbeiderskinderen (schoolscore > 1.15) en scholen met relatief weinig arbeiderskinderen (schoolscore < 1.15).

– In hoeverre bestaan er verschillen binnen leeftijds-groepen. Is het inderdaad zo, zoals we reeds stelden, dat er al bij het begin van de basisschool tussen leerlingen grote verschillen bestaan en worden deze gedurende het kleuteronderwijs opgeheven of worden de verschillen alleen maar vergroot: nivelleert de school de verschillen of treedt het zg. Mattheus-effect op? (Stanovich, 1986).

3 *Methode*

3.1 *Proefpersonen*

Bij het onderzoek waren 723 kleuters betrokken in de leeftijd van 4-6, 5 jaar, verdeeld over vijf leeftijdsgroepen van een half jaar. De verdeling van de leerlingen is vermeld in Tabel 1.

Tabel 1 *Steekproef verdeeld naar 5 leeftijdsgroepen, absoluut en in percentages*

Leeftijdsgroepen				Absoluut	%
1	4	- 4,5	jaar	46	6.4
2	4,5	- 5	jaar	97	13.4
3	5	- 5,5	jaar	294	40.7
4	5,5	- 6	jaar	176	24.3
5	6	- 6,5	jaar	110	15.2

Het betrof een landelijke steekproef waarbij per school 12 kinderen van een bepaalde leeftijdsgroep a-select werden gekozen.

Om vast te kunnen stellen of de resultaten van het onderzoek geldig zijn voor de leerlingen van groep 1 en 2 in de Nederlandse basisscholen is nagegaan in hoeverre de steekproef representatief was voor de Nederlandse populatie. Deze bleek in redelijke mate representatief te zijn. Zie voor een uiteenzetting hierover Van Kuyk en Kamphuis, (1990).

4 *Toetsprogramma*

Het toetsprogramma 'Ordenen' dat bij het onderzoek gebruikt is, bestaat uit elf onderdelen, waarvan drie bestaan uit probleemoplossen. De onderdelen zijn: 1. het kennen van kleuren; 2. het kennen van vormen; 3. het kennen van groottebegrippen ; 4. classificeren; 5. seriëren; 6. vergelijken; 7. tellen; 8. getallen; 9. probleemoplossen (PO) classificeren; 10. probleemoplossen seriëren; 11. probleemoplossen vergelijken en tellen.
In Figuur 1 worden de onderdelen vermeld met een korte omschrijving van de (deel)-vaardigheid die in het betreffende onderdeel wordt getoetst.

Ieder onderdeel is geoperationaliseerd in een toets met 15-30 opgaven, waarin steeds een toename in moeilijkheidsgraad van de items is nagestreefd. Per toets zijn twee parallelversies ontwikkeld die individueel worden afgenomen en op één uitzondering na dichotoom gescoord worden.

Het probleemoplossen verschilde in die zin van de andere onderdelen, dat het kind hier zelf de bedoelde kenmerken moest zoeken, terwijl ze bij de andere onderdelen werden gegeven. Het betrof probleempjes waarop één oplossing past.

5 *Het onderzoek*

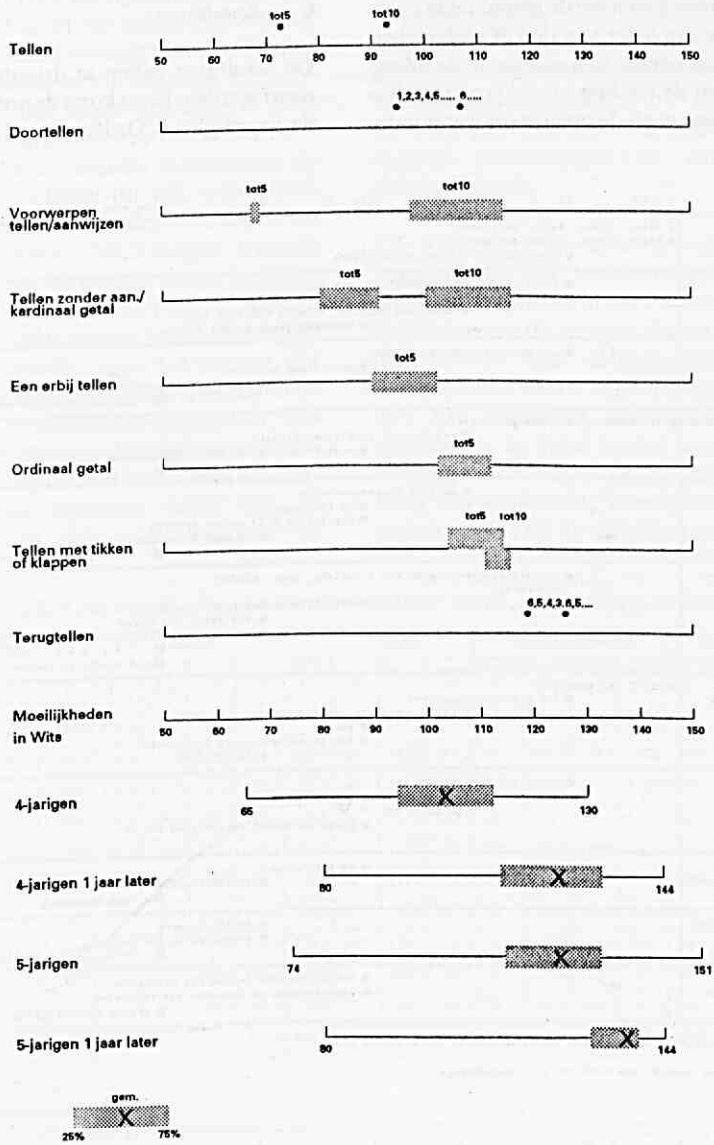
Het onderzoek werd uitgevoerd bij de vijf genoemde leeftijdsgroepen kleuters. In prin-

	Omschrijving vaardigheid
Kleur	: kennis van kleuren en kleurnuances
Vorm	: zien van overeenkomsten en verschillen bij eenvoudige en complexe figuren
Grootte	: hanteren van grootte-begrippen in stellende, vergrotende en overtreffende trap
Classificeren	: groeperen van objecten op een of meer kenmerken
Seriëren	: bepalen van de volgorde van objecten, zonder en met irrelevante kenmerken
Vergelijken	: bepalen van hoeveelheden t/m 10 in termen van meer, minder, evenveel, zonder en met irrelevante kenmerken
Tellen	: bepalen van aantallen t/m 10 met verschillende soorten telhandelingen
Getallen	: hanteren van getallen t/m 10 en de relatie tussen getal en aantal
PO. Classificeren	: convergente classificatieprobleempjes oplossen
PO. Seriëren	: convergente seriatieprobleempjes oplossen
PO. Tellen/vergelijken	: convergente tel/vergelijkprobleempjes oplossen

Figuur 1 *Overzicht van vaardigheden*

cipe zouden alle leeftijdsgroepen met elkaar vergeleken moeten worden op alle onderdelen van het programma. Alle kinderen zouden dus alle toetsen moeten maken. Omdat deze opzet onmogelijk kon worden uitgevoerd, zeker niet voor jongste kleuters, werd een zodanig design gekozen, dat het mogelijk was alle leeftijdsgroepen met elkaar te vergelijken op alle toetsen, zonder dat alle kinderen alle toetsen maakten. Voor een beschrijving van het design verwijzen we naar Van Kuyk en Kamphuis (1990).

De verwerking van de onderzoeksgegevens vereist hier enige aandacht. Er werd geen gebruik gemaakt van de klassieke testtheorie, maar van de itemresponsetheorie, het één-parameter logistische testmodel van Rasch (Hambleton & Swaminathan, 1983). Op grond van deze theorie is het, met inachtneming van een aantal voorwaarden, mogelijk de opgaven van een toets qua moeilijkheid op een schaal af te beelden. Op deze wijze ontstaat een 'maatstok' waarop de vaardigheid van een leerling of een groep leerlingen kan worden afgelezen. Itemmoeilijkheid en vaardigheid worden in dezelfde maateenheden uitgedrukt (zgn. WITS-waarden, positieve getallen tussen 50 en 150). Doordat de opgaven in volgorde van



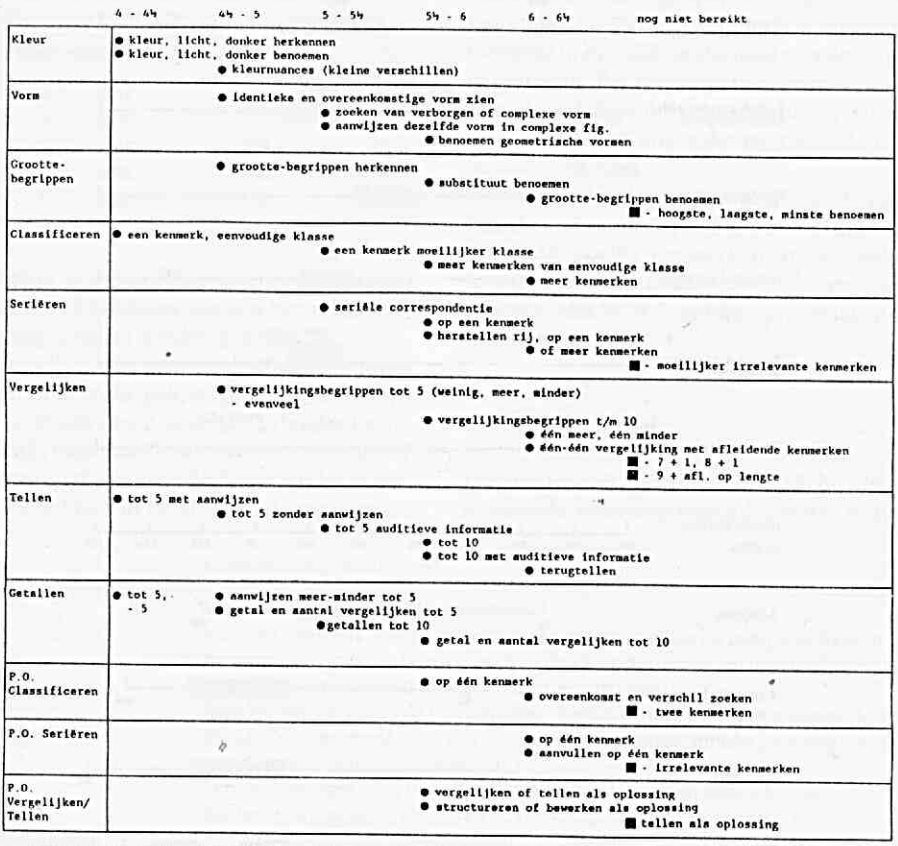
Figuur 2 *Schaal Tellen*

moeilijkheid zijn afgebeeld, is het niet alleen mogelijk de score van een leerling weer te geven, maar ook de mate waarin deze een vaardigheid heeft ontwikkeld of beheerst.

In dit geval hebben we ieder onderdeel dat bestaat uit een (deel)vaardigheid van 'Ordenen' gesplitst in een aantal subvaardigheden die als didactische eenheden kunnen worden opgevat. Hieronder verstaan we eenheden waarmee de leerkracht in het onderwijs kan handelen. Als voorbeeld geven we hier de schaal van het onderdeel 'Tellen' weer, waarin we een aantal subvaardigheden hebben onderscheiden (bijvoorbeeld: tellen tot vijf, mèt of zonder aanwijzen) die alle op dezelfde schaal passen. Tevens geven we de gemiddelde vaardigheid aan van ieder van de vijf onderscheiden leeftijdsgroepen. Een positie op de schaal betekent, dat de leerling of een groep leerlingen 50 % kans heeft de opgave (of wat minder specifiek de subvaardigheid) die met het vaardigheidsgetal overeenkomt goed te maken (vergelijkbaar met de 'maatstok' die bij hoogspringen wordt gebruikt. Als de lat ligt op het vaardigheidsniveau van de springer, heeft deze 50 % kans dat de lat blijft liggen). Alle onderdelen zijn op deze wijze geschaald en bleken ook op een schaal te passen (zie verder Van Kuyk & Kamphuis, 1990). Door op alle schalen de vaardigheid van de vijf leeftijdsgroepen af te beelden kunnen voor alle onderdelen 'mijlpalen' in de ontwikkeling of fasen in het onderwijsproces worden bepaald.

6 *Resultaten*

De resultaten zullen in drie delen gepresenteerd worden. Eerst komt de ontwikkeling van de vaardigheid 'Ordenen' gedurende de vijf



Figuur 3 *Overzicht van de vaardigheden van ordenen van verschillende leeftijdsgroepen 4-6½*

leeftijdsperioden aan bod, de zgn. 'mijlpalen' in de ontwikkeling, vervolgens wordt nagegaan in hoeverre er verschillen bestaan tussen leerlingen wat hun achtergrondkenmerken betreft en ten slotte wordt gekeken of er binnen leeftijdsgroepen verschillen bestaan.

6.1 *'Mijlpalen' in de ontwikkeling*

Omdat de plaats in dit artikel het niet toelaat volstaan we in Figuur 3 met de weergave van een samenvatting van de gemiddelde vaardigheid per leeftijdsgroep die we opvatten als 'mijlpalen': belangrijke ontwikkelingsmomenten die voor de leerkracht een handvat kunnen bieden voor het didactisch handelen.

Uit Figuur 3 blijkt, dat enkele deelvaardigheden van 'Ordenen' vroeger tot ontwikkeling komen dan andere, terwijl sommige deelvaardigheden meer tijd vergen dan andere. Zo is kennis van kleuren en vormen al bij 4-5-jarigen ontwikkeld, evenals classificeren op één kenmerk, en tellen tot vijf, terwijl probleemoplossen pas bij oudere leerlingen tot ontwikkeling komt: 5-5,5-jarigen (toetsen alleen afgenomen bij de drie oudste leeftijdsgroepen). Sommige 'mijlpalen' worden gemiddeld genomen tijdens het kleuteronderwijs nog niet bereikt, zoals het kennen van enkele groottebegrippen in de overtreffende trap, seriëren op irrelevante kenmerken, vergelijken op irrelevante kenmerken en probleemoplossen op meer of irrelevante kenmerken. Over het algemeen blijkt dat de deelvaardigheden zich geleidelijk over de kleuterleeftijd ontwikkelen, eerder gelijktijdig dan na elkaar. Kennis van kleur, vorm en grootte ontwikkelt zich slechts in beperkte mate vooraf aan classificeren, seriëren en vergelijken, waarvoor deze kennis nodig is. Probleemoplossen komt wel later in de ontwikkeling dan de andere vaardigheden waarvan verondersteld mag worden dat de kenmerken bekend moeten zijn voordat ze gezocht kunnen worden. Binnen de onderdelen zijn op subvaardigheidsniveau duidelijke ontwikkelingsverschillen te constateren. Globaal gesproken kan men zeggen dat eerst het herkennen van begrippen als 'mijlpaal' geldt en daarna het benoemen; met betrekking tot classificeren, seriëren en vergelijken eerst handelingen met één kenmerk en daarna met meer of irrelevante kenmerken; met betrekking tot de vaardigheid met aantallen tot 5 en daarna tot 10. Voor probleemoplossen geldt: eerst eenvoudige probleempjes en daarna meer gecompliceerde.

6.2 *Verschillen door achtergrondkenmerken*

Met behulp van variantie-analyse is per onderdeel nagegaan in hoeverre er verschillen bestaan tussen de groepen die we onderscheiden hebben naar achtergrondkenmerken: leeftijd, sekse, aantal maanden minder school dan het maximum aantal maanden, en schoolscore. In Tabel 2 is het totale overzicht weergegeven. Iedere onderscheiden groep kreeg een gemiddelde vaardigheidsscore uitgedrukt in WITS. In de kolom gemiddelde is

Tabel 2 *Gemiddelden qua leeftijd, sekse, aantal maanden minder school, en schoolscore per onderdeel*

	Gem.[1]	Leeftijd						Sekse			Aantal mnd school				Schoolscore		
		P	L1	L2	L3	L4	L5	P	j	m	P	0	−1	−2	P	< 1.15	− 1.15
Kleur	122.4	.00**	114.1	121.8	128.9			.00**	119.8	125.3	.05*	123.1	122.9	115.0	0.2*	123.4	117.7
Vorm	113.9	.00**	103.3	111.2	114.8	123.3	120.8	.02*	112.7	115.2	.51	114.0	114.5	111.6	.01**	114.9	111.1
Grootte	116.1	.00**	104.9	112.0	117.6	123.5	128.4	.10	115.1	117.0	.82	116.0	117.0	115.2	.00**	116.9	111.3
Classificeren	112.4	.00**	99.0	104.9	108.5	118.6	120.5	.09	111.6	113.1	.43	112.7	111.3	111.1	.00**	114.2	105.1
Seriëren	105.9	.00**	95.2	98.3	102.3	112.4	111.4	.06	105.0	106.7	.39	105.9	107.0	104.1	.00**	107.6	99.8
Vergelijken	112.3	.00**	94.0	100.2	112.1	116.0	118.6	.09	111.1	113.3	.83	112.4	112.7	110.8	.00**	113.4	107.9
Tellen	120.6	.00**	93.5	103.3	116.6	130.6	132.7	.01*	118.9	122.1	.13	121.2	117.7	119.4	.00**	122.1	115.3
Getallen	117.5	.00**	95.4	104.3	113.6	124.6	129.3	.00**	114.4	120.0	.23	117.5	115.5	120.5	.00**	119.6	111.6
PO Clas.	105.4	.00**			100.6	106.6	107.6	.42	105.9	105.0	.32	105.4	106.6	103.2	.00**	108.2	98.8
PO Seriëren	99.2	.00**			95.5	101.0	99.6	.32	98.7	99.6	.39	99.1	100.5	97.9	.00**	100.7	95.3
PO Verg/Tell	109.4	.00**			102.1	112.0	111.8	.66	109.1	109.7	.77	109.6	108.6	109.3	.00**	111.2	105.8

* 5% significantieniveau
** 1% significantieniveau
1) Bij de intelligentietests gelden de gewone gemiddelden en niet de WITS-waarden.

het algemeen gemiddelde aangegeven van de totale groep per onderdeel.
P geeft het significantieniveau.

Uit de gemiddelden over de groepen heen blijkt dat de onderdelen kleur, tellen en getallen de gemakkelijkste deelvaardigheden zijn en seriëren, probleemoplossen classificeren en probleemoplossen seriëren de moeilijkste. Hieronder geven we de verschillen per achtergrondvariabele weer.

6.2.1 *Leeftijd*
Tussen de vijf leeftijdsgroepen blijken significante verschillen te bestaan, zoals te verwachten was. Tussen leeftijdsgroepen 1 en 4 zijn de verschillen het grootst, tussen leeftijdsgroep 4 en 5 zijn de verschillen gering en bij enkele onderdelen in de tegenovergestelde richting. Het lijkt erop dat groep 5 zich niet verder ontwikkelt of niets meer bijleert. Uit de waarden per leeftijdsgroep blijkt (zie ook de gemiddelden curve in Bijlage 1) dat kleur, grootte, tellen en getallen een sneller ontwikkelingsverloop te zien geven dan vorm, vergelijken, classificeren en seriëren; een nog trager ontwikkelingsverloop is te zien bij de vaardigheden probleemoplossen classificeren en probleemoplossen seriëren. Dat geldt niet voor probleemoplossen vergelijken/tellen.

6.2.2 *Sekse*
Als we de sekseverschillen in ogenschouw nemen, zien we dat meisjes over vrijwel alle vaardigheden hogere prestaties leveren dan jongens. Significant zijn de verschillen bij kleur ($p < .01$), vorm ($p < .05$) tellen ($p < .01$) en getallen ($p < .01$). Bij probleemoplossen zijn de verschillen tussen jongens en meisjes te verwaarlozen.

De tendens dat meisjes over de hele linie hoger scoren dan jongens is opmerkelijk omdat uit gegevens van de PPON-rekenpeiling medio basisschool blijkt, dat jongens in belangrijke mate hoger scoren dan meisjes, behalve bij de eenvoudige opgaven (Wijnstra, 1988). Deze tendens dat meisjes op eenvoudige vaardigheden hoger scoren dan jongens wordt in ons onderzoek wel bevestigd. De tendens van medio basisschool dat jongens significant hoger scoren dan meisjes wordt in de rekenpeiling eind basisschool nog versterkt (Wijnstra, 1988).

6.2.3 *Aantal maanden minder onderwijs*
Uit de onderzoeksgegevens blijkt dat vrijwel alle kinderen onmiddellijk na het bereiken van de vierjarige leeftijd naar de basisschool gaan. Slechts weinig kinderen gaan 1, 2 of meer maanden later naar school. Van de 700 kinderen waarvan bekend is wanneer ze naar school zijn gegaan, gaan 642 meteen of na een maand naar school. 58 Kinderen gaan twee maanden later of meer. Dat is slechts 8%. De vraag of het aantal maanden dat een kind minder naar school gaat dan gezien zijn leeftijd mogelijk is ertoe doet, lijkt daarom niet beantwoord te kunnen worden.

Uit Tabel 2 blijkt dat kinderen die twee maanden later of nog later naar school gaan in de meeste gevallen (behalve bij tellen, getallen en probleemoplossen vergelijken/tellen) lagere prestaties leveren dan hun leeftijdgenoten die meteen bij het vier jaar worden naar school gaan, maar de verschillen zijn alleen bij het onderdeel kleur ($p < .05$) significant.

6.2.4 *Schoolscore*
Veel duidelijker zijn de verschillen tussen kinderen van scholen met een schoolscore < 1.15

Tabel 3 *Verschillen tussen leeftijdsgroepen 2SD boven en beneden $\bar{x}$ in Wits voor het onderdeel tellen*

Leeftijds-groepen	Wits			
	- 2SD	$\bar{x}$	+ 2SD	Verschil
Tellen				
1: 4 - 4½	47.49	95.18	142.88	95.39
2: 4½ - 5	74.76	105.12	135.48	60.72
3: 5 - 5½	86.98	116.50	146.02	59.04
4: 5½ - 6	104.99	130.41	155.84	50.85
5: 6 - 6½	110.00	131.60	153.19	43.19
Vooruit	62.51	36.42	10.31	

en van scholen met een schoolscore >1.15, scholen met relatief veel arbeiderskinderen. Uit Tabel 2 blijkt dat kinderen van eerstgenoemde scholen op alle onderdelen significant hoger scoren ($p < .01$; bij kleur, $p < .05$) dan kinderen van laatstgenoemde scholen. Deze verschillen blijven in de loop van de basisschool bestaan, zij het dat in de genoemde rekenpeiling van PPON (Wijnstra, 1988) de verschillen tussen de gewichten (1.00-1.70-1.90) aanmerkelijk groter zijn. Hierbij zijn ook veel allochtone kinderen betrokken. Aan het eind van de basisschool zijn de verschillen nog significant, maar de verschillen zijn enigszins kleiner geworden.

6.3 *Verschillen tussen hoog- en laagscorenden*

Om na te gaan in hoeverre er verschillen bestaan binnen leeftijdsgroepen of deze verschillen al bij het begin van het kleuteronderwijs bestaan en of deze verschillen gedurende het onderwijs toe- of afnemen, werd per leeftijdsgroep onderscheid gemaakt tussen hoog- en laagscorenden. Met name zijn we geïnteresseerd in de wijze waarop de laagscorenden functioneren met betrekking tot de verschillende onderdelen. Hoog- en laagscorenden worden als volgt gedefinieerd: hoogscorenden zijn leerlingen die een score hebben twee of meer standaarddeviaties hoger dan het gemiddelde; laagscorenden hebben een score twee of meer standaarddeviaties lager dan het gemiddelde. Ongeveer 95% van de leerlingen valt buiten deze range; van de overblijvende 5% scoort 2,5% hoger (hoogscorenden) en 2,5% lager (laagscorenden) dan deze leerlingen. In Tabel 3 is dit voor het onderdeel tellen aangegeven.

In Bijlage 1 is voor alle onderdelen het ontwikkelingsverloop van de hoog- en laagscorenden aangegeven, met daartussen het gemiddelde ontwikkelingsverloop. Hieruit blijkt dat de laagscorenden in het algemeen meer vooruitgang boeken dan de hoogscorenden. Dat is mede het gevolg van een plafondeffect. Indien voor hoogscorenden een hoger niveau gemeten kon worden dan het onderhavige instrumentarium aangeeft, zouden hoogscorenden meer vooruitgang hebben kunnen boeken.

Als met het instrumentarium het vigerende curriculum wordt bestreken, dan is er voor hoogscorenden vanaf het begin van het kleuteronderwijs weinig bij te leren, behalve voor vorm, grootte, seriëren en enigszins voor tellen en getallen. Dit is ook duidelijk te zien in Tabel 4 waar de verschillen in WITS-waarden zijn weergegeven.

Laagscorenden maken sterke vorderingen, behalve bij vorm en seriëren. Met name bij dit onderdeel lijken laagscorenden gedurende de hele kleuterperiode geen vooruitgang te boeken. Hier is mogelijk sprake van een bodemeffect. Dat wil zeggen dat deze groep leerlingen nog niet toe is aan betreffende vaardigheid. Bij probleemoplossen gaan noch de hoog-, noch de laagscorenden erg vooruit. Het lijkt erop dat de laagscorenden een deel van de achterstand die ze in het begin hebben, inlopen gedurende de eerste periode van het kleuteronderwijs: de armen lijken rijker te worden (zie bijvoorbeeld tellen), maar deze nivellering zet niet door in de laatste periode. Hierbij moet echter in het oog gehouden worden dat het maximaal haalbare voor de hoogscorenden niet gemeten is. Waar geen plafondeffect aanwezig is maar een bodemeffect zoals bij seriëren, worden de verschillen tussen hoog- en laagscorenden zelfs groter.

Tabel 4 *Vooruitgang van laag- en hoogscorenden in twee jaar onderwijs en het verschil tussen hoogscorende 4-jarigen en laagscorende 6-jarigen in WITS*

	Lage score vooruit	Hoge score vooruit	Verschil 4 hoog - 6 laag
Kleur	22	7	37
Vorm	12	19	22
Grootte	35	12	25
Classificeren	31	6	8
Seriëren	6	22	24
Vergelijken	34	10	37
Tellen	63	10	33
Getallen	43	21	35
PO Classificeren	8	7	34
PO Seriëren	3	3	28
PO Vergelijken/ Tellen	6	9	24
Totaal	263	126	337

Bij het interpreteren van de getallen dient te worden opgemerkt, dat bij kleur slechts leeftijdsgroep 1-3 en bij probleemoplossen slechts leeftijdsgroep 3-5 betrokken zijn.

Kijken we naar de verschillen tussen hoogscorende vierjarigen en laagscorende zesjarigen, dan is het opmerkelijk dat de eersten bij

het begin van het kleuteronderwijs al een score hebben, die door laagscorende zesjarigen in geen enkel geval wordt gehaald aan het einde van het kleuteronderwijs. Tussen hoogscorende vierjarigen die nog geen onderwijs hebben genoten en laagscorende zesjarigen die 2-2½ jaar kleuteronderwijs hebben gehad blijkt een aanzienlijk verschil, ten nadele van de laagscorende zesjarigen.

7 *Conclusies*

De antwoorden op de vraag wat kleuters van 4-6 jaar presteren op het gebied van 'Ordenen' gedurende de eerste twee klassen van de basisschool geven aanleiding tot de volgende conclusies.

Op grond van de verzamelde onderzoeksgegevens is het mogelijk een aantal 'mijlpalen' in de ontwikkeling vast te stellen. Er kan duidelijk worden aangegeven welke prestaties jongste en oudste kleuters leveren en welke prestaties ze nog niet kunnen leveren aan het eind van het kleuteronderwijs. De onderdelen van 'Ordenen' ontwikkelen zich voor een belangrijk deel gelijktijdig gedurende de periode 4,5-6 jaar. Kennis van kleur gaat daaraan voor een deel vooraf, probleemoplossen volgt voor een groot deel.

Binnen de vaardigheden zijn enkele algemene 'mijlpalen' te onderkennen: herkennen van begrippen gaat vooraf aan het benoemen ervan; handelingen met één kenmerk gaan vooraf aan handelingen met meer of irrelevante kenmerken; aantallen tot 5 gaan vooraf aan aantallen tot 10 en eenvoudige probleempjes gaan vooraf aan complexe probleempjes. Vooral de toename van complexiteit bepaalt binnen de onderdelen de 'mijlpalen' bij 'Ordenen'

Met betrekking tot de achtergrondkenmerken zijn er enkele duidelijke verschillen te constateren: leeftijd blijkt een bepalende factor te zijn. Door toename van de leeftijd worden de scores van de leerlingen hoger, behalve bij de hoogste leeftijdsgroepen. De in het voorgaande genoemde 'mijlpalen', die gebaseerd zijn op leeftijden berusten dus niet op toeval, maar op werkelijke verschillen. Sekse blijkt op enkele onderdelen verschillen te geven: kleur, vorm, tellen en getallen. Vooral bij de onderdelen waar eenvoudige handelingen van de leerlingen verwacht worden scoren meisjes hoger dan jongens. Dat geldt niet voor complexer handelingen als probleemoplossen, waarop jongens overigens niet hoger scoren dan meisjes, zoals dat later wel het geval is blijkens PPON-rekenonderzoek (Wijnstra, 1988). Het aantal maanden minder onderwijs dan maximaal mogelijk is, blijkt geen significant verschil te maken. Het aantal leerlingen dat minder naar school gaat is zeer klein en het aantal dat meer dan twee maanden minder gaat is gering. Of een kind op een school zit met een schoolscore van < 1.15 of > 1.15 maakt een groot verschil. Dat geldt voor alle onderdelen van 'Ordenen'. Kinderen die op een school zitten met een schoolscore van > 1.15 zijn ten opzichte van kinderen die op een school zitten met een schoolscore van < 1.15 in het nadeel.

De verschillen tussen leerlingen binnen dezelfde leeftijdsgroepen zijn groot bij het begin van het basisonderwijs en blijven groot gedurende het hele kleuteronderwijs. De verschillen lijken op een aantal onderdelen verkleind te worden, maar op andere worden ze vergroot. Vooral als we kijken naar de verschillen tussen hoogscorende vierjarigen en laagscorende zesjarigen, kunnen we concluderen dat de eersten niet veel kunnen bijleren en dat de laagscorenden weliswaar hebben bijgeleerd maar dat de verschillen aan het eind van twee jaar met de beginsituatie van de vierjarige hoogscorenden nog aanzienlijk zijn. Twee jaar kleuteronderwijs blijkt niet voldoende om deze kloof te overbruggen; in tegendeel, na twee jaar zijn de verschillen tussen hoog- en laagscorenden nog aanzienlijk en bij sommige vaardigheden zoals vorm en seriëren zelfs toegenomen tussen begin en eind van het kleuteronderwijs.

Doordat noch hoog- noch laagscorenden vooruitgang boeken op het gebied van probleemoplossen lijken deze onderdelen braakliggend terrein in het kleuteronderwijs.

8 *Discussie*

Op de vraag wat feitelijk het resultaat is van twee jaar kleuteronderwijs op het gebied van 'Ordenen' dient een genuanceerd antwoord gegeven te worden. We hebben gesteld dat we bij het vaststellen van resultaten in de eerste plaats geïnteresseerd waren in de resultaten van kinderen die ten opzichte van hun leeftijdgenoten achterstanden vertonen en op grond daarvan speciale hulp nodig hebben. Welnu,

aan de hand van de 'mijlpalen' is duidelijk vast te stellen dat er kinderen zijn die ten opzichte van leeftijdgenoten achterstanden vertonen. Maar dat niet alleen. Op basis van het verzamelde materiaal is het eveneens mogelijk doelstellingen te formuleren die op een bepaald tijdstip kunnen worden nagestreefd. Op grond hiervan kan niet alleen vergeleken worden met een normgroep, maar ook met een inhoudelijk criterium. Hierdoor kan een doelgerichte hulpverlening worden opgezet (zie ook Figuur 3). Op deze wijze worden de 'mijlpalen' diagnostisch interpreteerbaar en didactisch hanteerbaar.

Uit de inleiding is wel duidelijk geworden dat voor leerlingen met achterstanden een omvattende gestructureerde aanpak wenselijk is, zo mogelijk individueel. In het verlengde van het toetsprogramma 'Ordenen' hebben wij een hulpprogramma ontwikkeld, waarin naast gerichte oefening in een één-éénsetting (leerkracht met kind) ook aandacht besteed wordt aan metacognitieve aspecten: anticiperen, gestructureerd uitvoeren en controleren van opdrachten. Een dergelijk programma dat geheel aansluit bij de vastgestelde 'mijlpalen' zou dus kunnen worden ingezet om kinderen met achterstanden hulp te bieden.

Maar de resultaten maken eveneens duidelijk dat de extra aandacht niet alleen gericht mag zijn op kinderen met achterstanden. Kinderen die de school binnenkomen met een hoog ontwikkelingsniveau, dat soms al bij het begin verder reikt dan het vigerende curriculum, verdienen eveneens speciale zorg. Ze kunnen zich met het huidige curriculum maar nauwelijks verder ontwikkelen. Als het hier gehanteerde curriculum, dat tot stand is gekomen op basis van een analyse van Nederlandse en buitenlandse praktijkprogramma's, het gangbare aanbod weerspiegelt, dan mogen we stellen dat voor deze leerlingen het aanbod te beperkt is. Voor deze groep kinderen zou het curriculum uitgebreid moeten worden. In dat verband sluit het advies van de ARBO, verwoord in Spelen en Jongleren (1990) om het aanbod van het onderwijs aan jonge kinderen te verbreden goed aan bij onze bevindingen. Jonge kinderen die al ver gevorderd zijn in hun ontwikkeling kunnen dan actief en zinvol blijven deelnemen aan het onderwijsproces. Dat het onderzoeksmateriaal zou kunnen dienen als basis voor curriculumontwikkeling gaat dan in zoverre op dat door de 'mijlpalen' het curriculum gestructureerd en gereguleerd kan worden, maar wat de omvang betreft te beperkt is om tegemoet te komen aan alle leerlingen, met name de hoger ontwikkelde kinderen. Het programma is dan ook in de eerste plaats bedoeld voor kinderen met achterstanden. Een longitudinale curriculumplanning zou hieraan tegemoet kunnen komen.

Wat de resultaten van dit onderzoek in ieder geval opleveren is de wenselijkheid van differentiatie en individualisering van het onderwijs aan jonge kinderen.

Als we kijken naar de verschillen tussen kinderen wat hun achtergrondkenmerken betreft is het zeer wenselijk aan scholen met veel arbeiderskinderen en allochtone kinderen extra faciliteiten te geven om deze kinderen betere kansen te bieden. De achterstanden die ze ten opzichte van kinderen van andere scholen vertonen op alle onderdelen, is tekenend.

De verschillen tussen jongens en meisjes leveren niet voldoende aanknopingspunten om voor jongens of meisjes speciale maatregelen te nemen of het zou moeten zijn om meisjes de voorsprong te laten behouden door ze te motiveren voor en te richten op complexer taken en problemen.

Binnen de klassen zo blijkt heel duidelijk uit de resultaten is klassikaal onderwijs niet wenselijk. De verschillen zijn vanaf het begin van het onderwijs groot en blijven voor een belangrijk deel bestaan. Speciale zorg zal dus voor zowel de leerlingen met achterstanden als voor leerlingen die al een hoog ontwikkelingsniveau hebben moeten worden gerealiseerd. Maar voor de groep die zich daar tussenin bevindt, de grootste groep, zal het onderwijs eveneens optimaal moeten verlopen. In dit verband bevelen we één van de adviezen van Slavin en Madden (1989) nog eens aan: regelmatige evaluatie. Hiervoor biedt het onderhavige toetsinstrumentarium aanknopingspunten. Op deze wijze kunnen aanbod en resultaten van onderwijs, ook tijdens de kleuterperiode op elkaar afgestemd worden.

Literatuur

ARBO, *Voorrang en achterstand.* Advies over een integraal beleid ter voorkoming en bestrijding van onderwijsachterstand. Zeist: ARBO, 1988.
ARBO *Spelen en Jongleren*, Zeist: ARBO, 1990.

Cunningham, A. E., *A Metacognitive Approach to Phonemic Awareness*. Berkeley: University of California, 1989.

Dumont, J. J. & J. F. W. Kok, *Curriculum schoolrijpheid*. 's Hertogenbosch: Malmberg, 1970.

Dronkers, J., Kanttekeningen bij het congres ter afsluiting van het Rotterdamse project Onderwijs en Sociaal milieu (1969-1988). *Pedagogische Studiën*, 1990, *67*, 4, 193-196.

Early Years Curriculum Group, *Early Cildhood Education*. Stoke-on-Trent: Trentlam books, 1989.

Elkind, D., *The hurried child*. Reading, MA: Addison Wesley, 1981.

Groenendaal, H. J., *Vroegtijdige hulpverlening aan zwakfunctionerende kleuters*. Verslag van een onderzoek. Proefschrift VU Amsterdam: 1978.

Hambleton, R. K. & H. Swaminathan, *Item response theory principles and applications*. Boston: Kluwer Nijhoff Publishing, 1983.

Kuyk, J. J. van, Vroegtijdige onderkenning in de kleuterschool. *Kind en adolescent*, 1982, *3*, 85-93.

Kuyk, J. J. van, *Ruimtelijke Oriëntatie*. Tilburg, Zwijsen, 1985.

Kuyk, J. J. van, Basisvorming in het kleuteronderwijs. In: J. J. van Kuyk (Ed.), *Basisvorming in de basisschool*. Tilburg: Zwijsen, 1987.

Kuyk, J. J. van, & F. F. H. Kamphuis *Ordenen. Valideringsonderzoek van een observatieprogramma op het gebied van de voorbereiding op het rekenen in de eerste groepen van de basisschool*. Arnhem: Cito, 1990.

Kuyk, J. J. van, Toetsontwikkeling en de praktijk van het onderwijs. In: A. M. C. van den Geld & A. F. D. van Veen (red.), *Pedagogiek in beweging*. Amsterdam: Uitgeverij Balans, 1990.

Leij, A. van der, Zorgverbreding in de kleuterfase. In: A. van der Ley, *Zorgverbreding bijdragen uit speciaal onderwijs aan basisonderwijs*. Nijkerk: Intro, 1985.

Ministerie van O & W, Concept beleidsnota *evaluatie WBO* over de periode 1985-1987. 's Gravenhage: Ministerie van O & W, 1989.

Parreren, C. F., *Ontwikkelend onderwijs*. Leuven: Acco, 1988.

Slavenburg, J. H., Resultaten stimulerings- en activeringsprogramma's zoals vermeld in internationale literatuur. In: J. H. Slavenburg & T. A. Peters (red.), *Het Project Onderwijs en Sociaal milieu: een eindbalans*. Project Onderwijs en Sociaal milieu. Rotterdam: Rotterdamse School Advies Dienst, 1989.

Slavenburg, J. H., Criteria en evaluatie, een moeilijke combinatie. *Pedagogische Studiën*, 1990, *67*, 4, 196-199.

Slavenburg, J. H., J. M. G. Leune, B. P. M. Creemers, T. A. Peters, Conclusies en aanbevelingen. In: J. H. Slavenburg & T. A. Peters (red.), *Het Project Onderwijs en Sociaal milieu: een eindbalans*. Project Onderwijs en Sociaal milieu. Rotterdam: Rotterdamse School Advies Dienst, 1989.

Slavin, R. E., & A. Madden, What works for students at risk: a research synthesis. *Educational Leadership*, 1989, *2*, 4-13.

Stallings, J. A. & D. Stipek, Research on early childhood and elementary school teaching programs. In: M. C. Wittrock, (Ed.), *Handbook of research on teaching* 3 ed., 727-753 New York: MacMillan, 1986.

Stanovich, K. E., Mathew-effect in reading. Some consequences of individual differences in the acquisition of litteracy. *Reading Research Quarterly*, 1986, *21*, 4, 360-408.

Stevens, L. M., Vroegtijdige onderkenning van in hun ontwikkeling bedreigde kleuters. In: K. Doornbos & L. M. Stevens, *De groei van het speciaal onderwijs*. 's Gravenhage: Staatsuitgeverij, 1987.

Weihart, D. P., Curriculum Quality in early Education. In: S. C. Kagan & E. F. Zigler (Eds.), *Early Schooling: The national debate*. New Heaven: Yale University Press, 1987.

Wijnstra J. M., *Balans van het rekenonderwijs in de basisschool*. Arnhem: Cito, 1988.

Curriculum vitae

J. J. van Kuyk studeerde onderwijskunde aan de Katholieke Universiteit Nijmegen. Daarvoor was hij in het onderwijs werkzaam. Na zijn afstuderen was hij projectleider van diverse projecten en publiceerde over diagnostiserend onderwijzen in de basisschool, in het bijzonder op het gebied van het kleuteronderwijs. Sinds 1979 is hij projectleider van het project 'Eerste fase basis- en speciaal onderwijs' van het Cito. Hij gaf colleges aan de MO-B opleiding Onderwijskunde van de Tilburgse Leergangen en is redacteur van de Onderwijskundige Brochurenreeks. In 1985 promoveerde hij aan de Katholieke Universiteit Brabant op een proefschrift 'Ruimtelijke Oriëntatie; ontwikkeling en validering van een observatie- en hulpprogramma voor kleuters in de basisschool' (promotor Prof.Dr. L. F. W. de Klerk). Hij ontwikkelt en onderzoekt programma's ten behoeve van jonge kinderen met ontwikkelingsachterstanden.

Adres: Cito, Instituut voor Toetsontwikkeling, Postbus 1034, 6801 MG Arnhem.

Manuscript aanvaard 10-9-'90

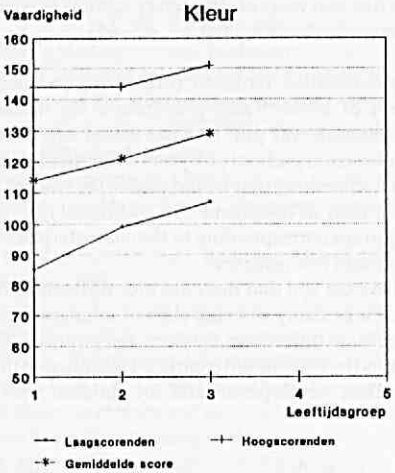
Summary

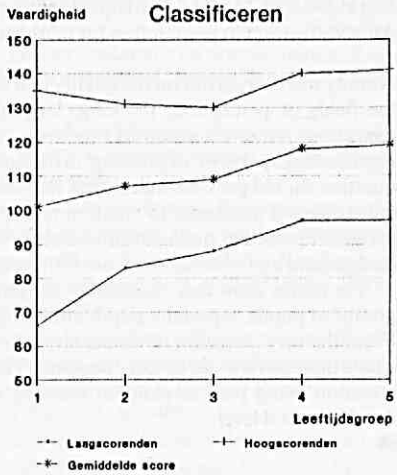
Kuyk, J. J. van. 'Are children capable of 'organizing' in the first two years of elementary school (4-6 years old)? An investigation in preparation for mathematics.' *Pedagogigsche Studiën*, 1990, *67*, 429-443.

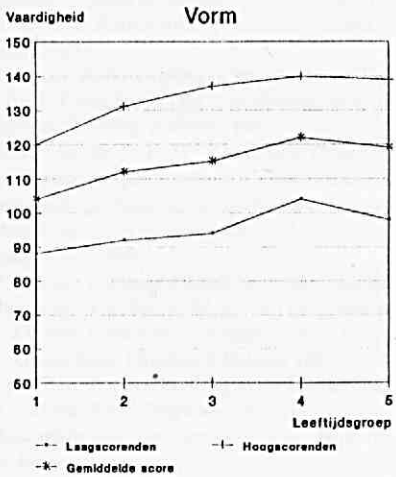
A study was carried out on the quality of a testing programme entitled 'Ordenen' (organizing), a basic skill in the fields of perception, thinking, language, the concept of number and preparation for mathematics education. As an extension to this study, the following question was put: to what extent can children be regarded as capable of 'organizing' in the first two years of elementary school (4-6 years old)? An answer to this question should provide 'milestones' marking development, something that would enable the teacher to offer more directed assistance to children not up to the mark in their development. An additional question was asked as to whether there are differences between various groups corresponding to the characteristics of their background and whether there are differences between high and low achievers.

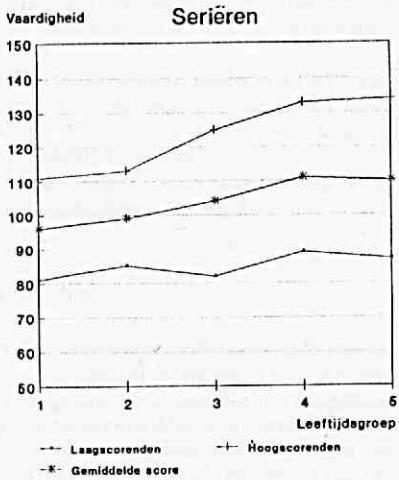
The results show that 'milestones' can indeed be shown to exist and that there are wide differences between groups of pupils, especially pupils attending schools with few or many working class or immigrant children. From the very inception of elementary education there are wide differences between age groups and on the whole these differences do not disappear. This discussion puts the case for differentiated teaching, with special attention being paid to children showing retardation in their development and for children with a high developmental level.

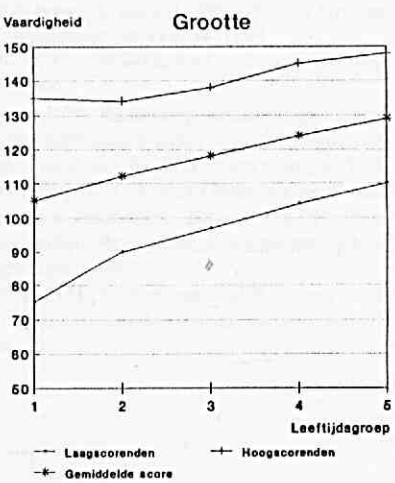
Bijlage 1

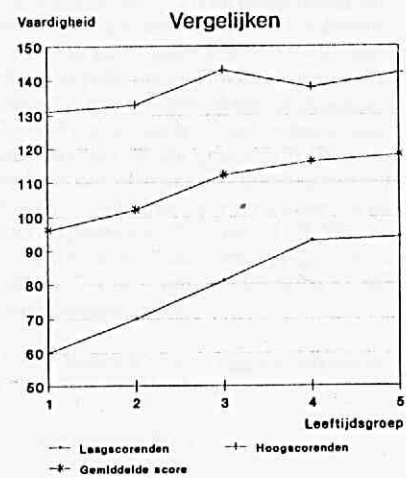

Vaardigheid
Kleur
160
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score

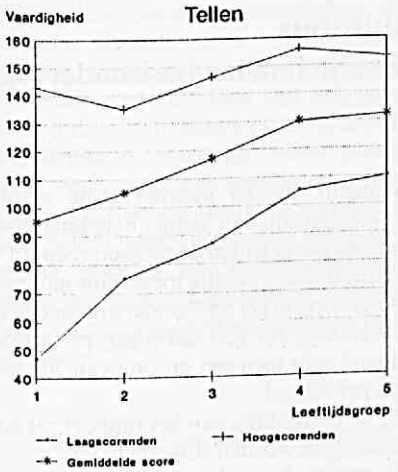

Vaardigheid
Classificeren
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score

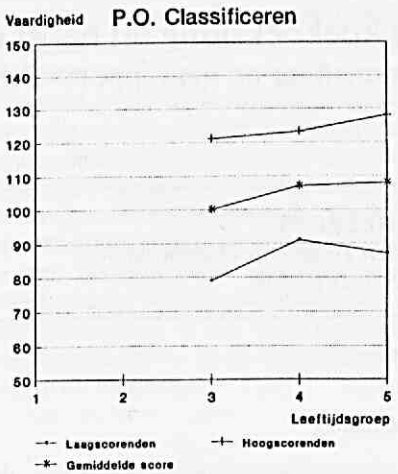

Vaardigheid
Vorm
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score

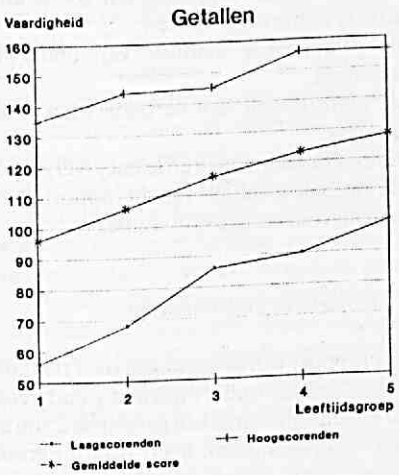

Vaardigheid
Seriëren
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score

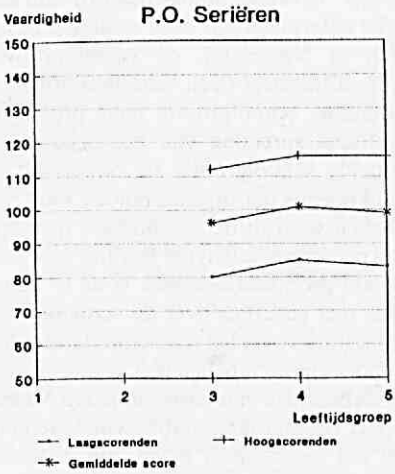

Vaardigheid
Grootte
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score

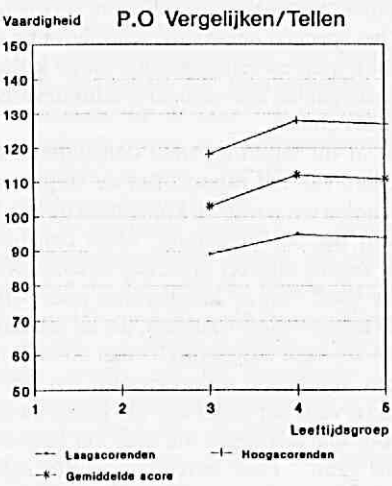

Vaardigheid
Vergelijken
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score


Vaardigheid
Tellen
160
150
140
130
120
110
100
90
80
70
60
50
40
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score


Vaardigheid
P.O. Classificeren
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score


Vaardigheid
Getallen
160
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score


Vaardigheid
P.O. Seriëren
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score


Vaardigheid
P.O Vergelijken/Tellen
150
140
130
120
110
100
90
80
70
60
50
1
2
3
4
5
Leeftijdsgroep
Laagscorenden
Hoogscorenden
Gemiddelde score

Discussie

Het koekoeksjong uit het ei van Columbus
Commentaar op het rapport "Schaal en kwaliteit in het basisonderwijs"

J. SCHEERENS
Faculteit Toegepaste Onderwijskunde, Universiteit Twente

1 *Inleiding*

Het rapport "Schaal en kwaliteit in het basisonderwijs" bevat het nodige beschrijvend statistische cijfermateriaal over aantallen basisscholen in Nederland, de verdeling over schoolgroottecategorieën, schoolgrootte naar denominatie, schoolgrootte naar provincie, geografische spreiding van basisscholen en gemiddelde schoolgrootte in Nederland in vergelijking met omringende landen. Ook zijn er tabellen waaruit de beschikbare formatie (aantal personeelsleden) per leerling voor verschillende groottecategorieën is af te lezen, tabellen met gegevens over de verdeling van schoolbesturen over het scholenbestand in het basisonderwijs en tabellen met gegevens over de bereikbaarheid van basisscholen in Nederland. Het belangrijkste statistische feitenmateriaal dat het rapport bevat zijn gegevens over de relatie tussen schoolgrootte en leerprestaties, schoolgrootte en doorverwijzing naar het speciaal onderwijs, schoolgrootte en produktiviteit en schoolgrootte en de kosten om een bepaalde hoeveelheid produktieverbetering te behalen.

Uit al dit cijfermateriaal concluderen de opstellers van het rapport dat de vergroting van scholen ten goede zal komen aan de kwaliteit van het basisonderwijs. Deze conclusie wordt kracht bijgezet door een beschouwing van de hand van de hoogleraren onderwijskunde Lagerweij en Claessen, die zij zelf aanduiden als "een redenering... met als stelling dat een voldoende schoolgrootte een voorwaarde is voor het onderwijs van goede kwaliteit voor alle leerlingen die naar het basisonderwijs gaan". Deze beschouwing dus moet aannemelijk maken *waarom* grote scholen beter zouden zijn dan kleine. Feitenmateriaal en beschouwing leiden de projectgroep tot de conclusie dat de wettelijk toegestane minimale schoolgrootte in het basisonderwijs moet worden verhoogd tot 250 leerlingen per school. Optimaal acht men een grootte van 500 leerlingen per school.

Bij de bespreking van het rapport zal achtereenvolgens worden stilgestaan bij:
- de empirische onderbouwing van de stelling dat schaalvergroting ten goede komt aan de onderwijskwaliteit;
- de gehanteerde definitie van onderwijskwaliteit;
- de plausibiliteit van de onderwijskundige analyse;
- de assumpties inzake efficiency-vergroting;
- de algehele kwaliteit van dit rapport als een beleidsvoorbereidende studie.

2 *Empirische onderbouwing*

In het rapport wordt gezwegen over tientallen voorbeelden van onderwijskundig onderzoek, waarbij schoolkenmerken gerelateerd zijn aan onderwijsopbrengsten, maar waarbij grootte nimmer naar voren is gekomen als een opbrengstvermeerderend schoolkenmerk. De ene, speciaal voor de projectgroep uitgevoerde, secundaire analyse van het SCP vormt, hoezeer in het rapport ook het tegendeel wordt gesuggereerd, geen uitzondering. Aannemende dat de bèta-gewichten in tabel 2.10 gestandaardiseerde bèta-gewichten zijn, verklaart de variabele schoolgrootte (bèta = 0,11) 1% van de variantie in schoolgemiddelden op de beide prestatietoetsen. Dit is een zeer, zeer zwak effect[1]. Het is nogal misleidend dat in het rapport wordt gesteld dat de bijdrage van schoolgrootte aan de totale variantie die gebonden wordt door de variabelen *samenstelling leerlingpopulatie*, *denominatie* en *schoolgrootte*, 'bescheiden', maar 'onmiskenbaar' is. Wanneer blijkt dat grotere scho-

Pedagogische Studiën 1990 (67) 444-449

len iets meer dan kleinere scholen doorverwijzen naar het speciaal onderwijs (een 'ongunstige' conclusie dus voor grote scholen), en het gestandaardiseerde regressiecoëfficiënt 0,08 bedraagt, spreken de opstellers wèl van een 'zeer zwakke relatie'. Men kan zich moeilijk aan de indruk onttrekken dat er op deze wijze versluierend en 'partijdig' wordt gerapporteerd over onderzoeksresultaten (zie ook Creemers & Hoeben, 1990). Als hier doorheen gezien wordt, blijkt de centrale stelling uit het rapport, waarbij een substantiële schoolgrootte als een noodzakelijke voorwaarde voor onderwijskwaliteit wordt gezien, op niets te berusten.

3 *De betekenis van onderwijskwaliteit*

"Pardon", zullen de opstellers van het rapport tegen deze kritiek inbrengen, "met onderwijskwaliteit bedoelen wij méér dan alleen maar onderwijsopbrengsten". In de bijlage van Claessen en Lagerweij wordt gesteld dat het begrip kwaliteit van het onderwijs niet "eenvoudig helder en precies is af te bakenen" en dat doen ze dan ook niet. In plaats daarvan onderscheiden zij kwaliteit 'in termen van' doelen en functies, inhouden, processen, personele en materiële voorzieningen en resultaten. Over deze 'aspecten van kwaliteit' wordt gesteld dat ze "nauw met elkaar samenhangen" (p. 89). Hoe die samenhang in elkaar zit wordt echter niet onthuld. Kennelijk hebben bepaalde kenmerken van processen, voorzieningen, etc. voor de auteurs een intrinsieke waarde. Het ontbreken van de notie dat de kwaliteit van processen, voorzieningen, etc. moet blijken uit de onderwijsresultaten leidt tot een verschijnsel waar de kwartaire sector toch al gemakkelijk aan lijdt, nl. 'goal displacement', het als doelen op zich gaan beschouwen van zaken die in feite middelen zijn om de werkelijke doelen van de organisatie te bereiken. In het onderwijs zijn die werkelijke doelen natuurlijk het bereiken van acceptabele onderwijsopbrengsten, hoe men die verder dan ook definieert. Het verschijnsel 'goal displacement' moet als een belangrijke bron van inefficiëntie van organisaties worden gezien. Reden om remedies, die aan dit verschijnsel doen denken in een analyse die juist op doelmatigheidsverbetering is gericht, met argusogen te bezien. De indruk van "goal displacement" wekt het rapport onder andere waar schoolvergroting (een term die naast schaalvergroting wordt gebruikt en waarvan men nooit zeker weet of er geen typefout is gemaakt) in feite als een middel wordt gezien om meer management mogelijk te maken. Normaal wordt in de organisatiekunde versterking en uitbreiding van het management onder meer als een *gevolg* van groei van de organisatie gezien. Dat "meer management", zoals Claessen en Lagerweij dat verder breed uitmeten met termen als beleidsvoerend vermogen, diagnostiserend vermogen, responsiviteit, etc. tot betere opbrengsten tegen minder kosten leidt, is een veronderstelling die weinig plausibel is. In de volgende paragraaf wordt dit toegelicht.

4 *De onderwijskundige analyse nader bezien*

In de bijlage van de hand van Claessen en Lagerweij, getiteld "Naar een grote school", wordt een groot aantal management- en organisatiekenmerken genoemd, welke de auteurs wenselijk achten. Vervolgens is dan steevast de conclusie dat om die wenselijke management- en organisatiekenmerken mogelijk te maken grotere scholen nodig zijn.

Hoewel een strakke onderbouwing van de stelling waarom de desbetreffende organisatiekenmerken wenselijk zijn achterwege blijft (en dat kan ook moeilijk gezien het ontbreken van een eenduidige kwaliteitsdefinitie) zijn er twee achterliggende redeneringen denkbaar:

a) de organisatiekenmerken in kwestie bevorderen een optimale inrichting van het onderwijsleerproces en daarmee indirect betere leerresultaten;

b) de organisatiekenmerken zijn noodzakelijk als reactie op veranderingen in de omgeving van de school; de van overheidswege in te voeren autonomievergroting van scholen is een voorbeeld van zo'n omgevingsverandering.

In het eerste geval worden management- en organisatiekenmerken gezien als randvoorwaarden die het primaire proces (het onderwijs dus) bevorderen en in het tweede geval is er sprake van condities (buffers) die het primaire proces van de organisatie moeten beschermen tegen ontwikkelingen in een veranderende omgeving. Beide benaderingen zijn bruikbare gezichtspunten bij het denken over effectieve schoolorganisaties. Een aldus opge-

bouwde zorgvuldige organisatie-analyse zal echter niet leiden tot de overdaad aan wenselijke management- en organisatiecondities die Claessen en Lagerweij oplepelen. In feite vegen zij lijsten met management- en organisatiekenmerken bij elkaar die uit geheel verschillende tradities binnen de onderwijskunde naar voren zijn gekomen, te weten het schooleffectiviteitsonderzoek, denkbeelden uit het aandachtsgebied van de 'planned change' en de schoolorganisatiemodellen van E.C.H. Marx.

Het schooleffectiviteitsonderzoek sluit het best aan bij de hierboven onder a) geformuleerde redenering. Verder biedt het schooleffectiviteitsonderzoek van de drie genoemde onderwijskundige tradities nog de meeste empirische evidentie, zij het dan dat bijvoorbeeld een concept als 'onderwijskundig leiderschap' veel minder empirisch bevestigd is dan uit de weergave van Claessen en Lagerweij blijkt. Overigens vormen instructiecondities op micro-niveau de kern van modellen, waarbij organistiekenmerken als bevorderende condities van effectieve instructie worden gezien (vgl. Scheerens & Creemers, 1989), en ligt de basis voor kwaliteitsverbetering in het onderwijs dus primair op het niveau van de inrichting van het onderwijsleerproces.

In het schoolveranderingsdenken zijn zaken als de zelfvernieuwende capaciteit of de veranderingscapaciteit van de school vaak de horizon van het denken. Ook in de hier aan de orde zijnde beschouwing worden dergelijke zaken kennelijk als intrinsieke kwaliteitsaspecten gezien. Het gevaar hiervan is dat er geïnvesteerd wordt in het versterken van de 'veranderbaarheid' van de school, zonder duidelijk voor ogen te hebben waartoe er veranderd moet worden, of de verandering ten goede komt aan de resultaten van het primaire proces, 'hoeveel aan veranderbaarheid' er nodig is om tegemoet te komen aan veranderende omgevingscondities en zonder in aanmerking te nemen dat aan een dergelijke investering ook kosten zijn verbonden; bijvoorbeeld in de zin van het onttrekken van energie aan het dagelijks werk op school.

Bij het derde referentiepunt in de analyse van Claessen en Lagerweij, de schoolorganisatiemodellen van E. C. H. Marx, is het individualiserend vermogen van de school in eerste instantie het criterium waarom het gaat. Bij toenemende individualisering zouden in toenemende mate geïntegreerde (en complexere) schoolorganisaties nodig zijn. In latere bijdragen vanuit deze benadering wordt meer aandacht gegeven aan de buffer-functie ten opzichte van externe ontwikkelingen; de term "beleidsvoerend vermogen" kan hiermee in verband worden gebracht (Marx, 1990). Eén van de problemen met deze modellen is dat er in feite met een onderwijskundig proceskenmerk (individualisering) als criterium wordt gewerkt. Hoe riskant dit is demonstreren onderzoeksuitkomsten waaruit blijkt dat geïndividualiseerd onderwijs, door een vermindering van de effectieve leer- en instructietijd, tot relatief slechtere leerprestaties leidt dan klassikaal onderwijs (vgl. bijv. Mortimore e.a., 1988).

Hoewel een deel van de door Claessen en Lagerweij als wenselijk aangeduide management- en organisatiekenmerken geplaatst kan worden in integrale schooleffectiviteitsmodellen (bijvoorbeeld het ontwikkelen van een prestatiegerichte 'missie' van de school en het evaluerend vermogen van de school) en een ander deel als noodzakelijke respons op externe ontwikkelingen kan worden gezien (bijvoorbeeld de ontwikkeling van vaardigheden om financieel beleid te voeren bij een grotere financiële autonomie) geldt dat:

- de kwaliteitsbevorderende werking van deze organisatiekenmerken nog goeddeels moet worden aangetoond;
- de betekenis van instructiecondities als meer directe beïnvloeders van onderwijsresultaten niet uit het oog moet worden verloren;
- het – voor zover de auteurs hebben gemeend een conceptueel ontwerp van de ideale school te moeten maken – te betreuren is dat er nauwelijks gebruik is gemaakt van de vele in de literatuur beschikbare integrale[2] conceptuele modellen van schooleffectiviteit;
- het gebrek aan een modelmatige conceptualisering van kwaliteitsbevorderende school- en instructiekenmerken leidt tot een overdaad aan aangeprezen schoolorganisatie- en managementcondities.

Vooral spaarzaamheid wordt node gemist in de analyse van Claessen en Lagerweij. Door de eclectische en structuurloze aanpak in het genereren van wenselijke organisatiekenmerken, door het niet hanteren van een rigoureuze output-georiënteerde kwaliteitsdefinitie en door het onderbelicht laten van bepaalde teleurstellende uitkomsten van empirisch

onderzoek wordt de indruk gewekt dat het geheel aan management- en organisatiecondities, dat in het rapport wordt opgesomd, 'meer management' dus, zonder meer goed zou zijn voor de kwaliteit en doelmatigheid van het onderwijs. Als deze denkwijze zou worden gevolgd bij de besteding van de miljoenen die men met schaalvergroting denkt te kunnen besparen, zou dit leiden tot vergroting van de 'overhead'-kosten in het onderwijs, meer geld dus voor *administration* en relatief minder geld voor verbetering van het primaire proces in het onderwijs. Op deze manier zou het middel wel eens erger kunnen blijken dan de kwaal en zou gevreesd moeten worden voor een uitdijende bureaucratie op het meso-niveau[3].

Van de onderliggende gedachtengang in de beschouwing van Claessen en Lagerweij: meer management bevordert onderwijskwaliteit en voor meer management zijn grotere scholen nodig, is de laatste schakel zo mogelijk nog dubieuzer dan de eerste. Bepaalde organisatorische wenselijkheden, zoals 'eenheid in visie' binnen de school lijken zelfs gemakkelijker bij kleine scholen gerealiseerd te kunnen worden, terwijl ook niet zo inzichtelijk is waarom kleine scholen meer te lijden zouden hebben van parttime aanstellingen dan grote scholen. Voor het overige geldt dat het, bij alle twijfels over de wenselijkheid van de hoeveelheid organisatie- en managementcondities die de auteurs aanvoeren, prematuur is om schaalvergroting als middel hiertoe te propageren onder het mom van kwaliteitsverbetering.

5 *Efficiëntievergroting door schaalvergroting?*

Het cijfermateriaal in het rapport lijkt overduidelijk aan te tonen dat grote scholen goedkoper zijn dan kleine scholen. Bij nadere inspectie van de tabellen blijkt dat het vooral de zeer kleine scholen zijn die relatief duur zijn. Het is merkwaardig dat er niet een uitvoeriger economische analyse is uitgevoerd. Wel bevat de bijlage van de hand van Claessen en Lagerweij een hoofdstuk met de titel "Een economische invalshoek", maar daarin wordt gesteld dat het door de complexiteit van het financieringssysteem moeilijk te becijferen is wat de opbrengst van substantiële schoolvergroting zal zijn, waarna de auteurs opmerken dat het hierbij 'naar hun indruk' om omvangrijke bedragen zal gaan. Over de analyses van het Sociaal en Cultureel Planbureau ten aanzien van de efficiencyvoordelen van schaalvergroting kan de lezer zich moeilijk een oordeel vellen, omdat de eigen rapportage daarover nog moet verschijnen en de meet- en rekenprocedures niet in het rapport van de projectgroep schoolgrootte zijn vermeld.

Elementen van een uitvoeriger economische analyse zouden kunnen zijn:

- het gebruik maken van bestaande literatuuroverzichten over efficiency-vergroting door schaalvergroting; een recent voorbeeld hiervan is het rapport van Kaiser en Vermeulen (1989), waaruit overigens blijkt dat er op grond van eerder onderzoek geen eenduidige conclusie inzake schaalvoordelen kan worden getrokken;
- het doorrekenen van alternatieven voor de hier voorgestelde maatregel om de minimale schoolgrootte in het basisonderwijs op een grootte van 250 leerlingen te bepalen; wat zou bijvoorbeeld de besparing zijn bij lagere minimumgrootten?
- het ontwerpen en doorrekenen van alternatieve scenario's om de gewenste concentratie van management en specialismen tot stand te brengen (onder meer met het oog op het terugbrengen van de toeloop naar het speciaal onderwijs); één zo'n scenario zou gebaseerd kunnen worden op de gedachte om netwerken van (eventueel kleine) scholen te vormen, waarbij bepaalde bestuurlijke management- en ondersteuningsfuncties geconcentreerd zouden kunnen worden;
- het in aanmerking nemen van de invoeringskosten van alternatieve varianten; hierbij valt onder meer te denken aan de kosten van nieuwe gebouwen[4], kapitaalvernietiging door het moeten afstoten van bestaande gebouwen, eventuele toename van vervoerskosten, zo mogelijk ook een inschatting van de kosten die gepaard gaan met de voorbereiding en het overleg inzake fusies van scholen e.d.

Bij het ontbreken van een dergelijke meer volledige economische analyse is het moeilijk om uit het rapport in kwestie definitieve conclusies te trekken over de kostenbesparing van schaalvergroting onder de aanname dat deze niet gepaard gaat met kwaliteitsverlies.

6 *Besluit: De kwaliteit van het rapport als beleidsvoorbereidende studie*

Het aan de orde zijnde rapport over schoolgrootte maakt onvoldoende gebruik van bestaande onderzoeksliteratuur op zowel het terrein van schooleffectiviteit als het terrein van efficiency-winst door schaalvergroting. Een grondige economische analyse ontbreekt goeddeels en de bijgevoegde onderwijskundige beschouwing is zwak qua onderbouwing en analyse. Voor zover 'eigen' onderzoeksresultaten worden gepresenteerd, wordt hierover op partijdige en versluierende wijze gerapporteerd. Ook in de presentatie van het cijfermateriaal wordt voortdurend op uitersten gekapitaliseerd: de allerkleinste worden met de allergrootste scholen vergeleken. Het rapport lijkt één grote poging om een vooraf vastgestelde conclusie te legitimeren en ontbeert het in aanmerking nemen van alternatieve maatregelen (deze laatste punten worden ook door Creemers & Hoeben, 1990, naar voren gebracht). Voor de stelling die uit het rapport met de meeste kracht opklinkt, namelijk dat schoolgrootte de onderwijskwaliteit bevordert, is het minste bewijs. Kennelijk wordt van deze formulering een grotere politieke overtuigingskracht verwacht, waar het misschien toch al voldoende was geweest om aan te tonen dat schaalvergroting, en de daarmee misschien verbonden kostenbesparing, niet ten koste gaat van het niveau van de onderwijsresultaten.

Drastische schaalvergroting, ook in het basisonderwijs, lijkt het ei van Columbus voor bewindslieden die graag iets aan het onderwijs zouden willen verbeteren, maar daar geen geld voor hebben. Bijstelling van de normen voor de minimale grootte van scholen is één van de weinige simpele mogelijkheden tot directe sturing door de overheid in het Nederlandse onderwijsbestel. En passant zou deze maatregel dan misschien gunstige voorwaarden kunnen creëren voor afbouw van de toeloop naar het speciaal onderwijs in het kader van het zogenoemde 'zorgverbredingsbeleid' en ook zou er, gezien de voorkeur welke wordt uitgesproken voor het onderbrengen van bepaalde ondersteuningsfuncties in de school, wellicht een basis worden gecreëerd om over te gaan tot 'rationalisering' van een deel van de verzorgingsstructuur, nl. de schoolbegeleidingsdiensten. Een gebrekkige beleidsanalyse die als surprise bovendien het koekoeksjong van een te uitbundige toename van schoolmanagementfuncties uit de hoge hoed tovert, zal in dit geval naar het zich laat aanzien eerder contraproduktief blijken ten opzichte van de respectabele doelstelling van doelmatigheidsverbetering in het onderwijs. Hopelijk zullen de voorstellen die uit deze gebrekkige analyse naar voren zijn gekomen niet door het beleid worden opgevolgd.

Noten

1. Wanneer, zoals meer gebruikelijk is, een dergelijk effect zou worden uitgedrukt in termen van gebonden variantie in de prestaties van individuele leerlingen zou het nog gedecimeerd worden.
2. Met 'integraal' wordt hier bedoeld dat deze modellen zowel schoolorganisatie als instructiecondities omvatten.
3. In een recente voordracht voor het Interuniversitair Centrum voor Onderwijsevaluatie schetste Herbert J. Walberg een afschrikwekkend beeld van de situatie in sommige Amerikaanse staten, waar het merendeel van onderwijsvernieuwingsgelden wordt opgeslokt door de 'administration'.
4. Het Informatie- en Adviescentrum Schoolaccommodatie schat de kosten van nieuwbouw op maar liefst *f* 1,7 miljard (Kooistra, 1990).

Literatuur

Creemers, B. P. M. & W. Th. J. G. Hoeben, Onderwijsonderzoek te kijk gezet; commentaar op het rapport "Schaal en kwaliteit in het basisonderwijs". *Tijdschrift voor Onderwijsresearch*, 1990, *15*, 234-241.

Kaiser, F. & C. J. A. J. Vermeulen, Het doelmatig functioneren van scholen. Enschede: Universiteit Twente, Onderzoekscentrum Toegepaste Onderwijskunde, 1989.

Kooistra, S., Schaalvergroting vergt miljarden. In: *NG Gemeentelijk Magazine*, 14 september 1990, p. 3.

Marx, E. C. H., *Schoolontwikkeling en milieugericht onderwijs*. Afscheidscollege. Alphen aan den Rijn: Samsom, Tjeenk Willink, 1990.

Mortimore, P. et al., *School matters: the junior Years*. Somerset: Open Books, 1988.

Projectgroep Schaalvergroting Basisonderwijs, *Schaal en kwaliteit in het basisonderwijs*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1990.

Scheerens, J. & B. P. M. Creemers, Conceptualizing school effectiveness. In: Developments in school effectiveness research (pp. 691-706). Special issue of the *International Journal of Educational Research*, 1989, vol. 13, no. 7 Oxford: Pergamon Press.

Curriculum vitae

J. Scheerens is als hoogleraar verbonden aan de Universiteit Twente, Faculteit Toegepaste Onderwijskunde. Zijn leeropdracht ligt op het terrein van onderwijsorganisatie en -management. Hij is momenteel voorzitter van het Interuniversitair Centrum voor Onderwijsevaluatie (ICO): een samenwerkingsverband voor 2de fase onderwijs van de Universiteit Twente en de Rijksuniversiteit Groningen en wetenschappelijk directeur van het Onderzoekscentrum Toegepaste Onderwijskunde (OCTO).

Adres: Universiteit Twente, Postbus 217, 7500 AE Enschede

Manuscript aanvaard 22-10-'90

Summary

Scheerens, J. 'School size and quality in education: a faulty policy-advice.' *Pedagogische Studiën*, 1990, *67*, 444-449.

The recent report on "school size and quality in education" is reviewed from the angle of empirical evidence, conceptual quality and policy relevance. The conclusion is that the report lacks an empirical basis and that the conceptual analysis is rather shaky. Moreover, counterproductive effects are expected, both in terms of neglected costs of increasing school size and additional costs for "more management" that are likewise neglected. The central thesis of the report: 'school size enhances educational quality' cannot be maintained.

De AERA Annual Meeting: Boston 1990

Twee werelden: Introductie (W. J. Nijhof)

Aangekomen in Boston vraagt een vrouwelijke douane-officier mij aan de hand van mijn inlichtingenformulier of ik "in het onderwijs" werkzaam ben. Ik bevestig dat. Machteloos slaat ze de handen ten hemel en roept "Come and help us". Ze probeert een 'Dutch professor' duidelijk te maken dat er wat aan de hand is met het Amerikaanse onderwijssysteem en dat dringend hulp van buiten geboden is, met name vanuit Nederland met ons schitterende onderwijs. Het deugt niet met de kwaliteit van het Amerikaanse onderwijssysteem roept ze, acht jaren Reagan waren desastreus. Zou iets van deze harde werkelijkheid doorklinken tijdens de AERA?

Op Paaszaterdag, de dag voor Pasen derhalve, tref ik een zaal met 80 volwassenen aan in een hotel halverwege tussen St. Louis en Carbondale bezig met een masters degree: volwasseneneducatie en universitair onderwijs. Een tandartsassistente, studerend voor haar masters in Education, komt na een Hollands college naar mij toe en laat mij de Wall Street Journal van februari zien, waarin voorspellingen staan over de werkgelegenheid. Deze zijn dramatisch: wie geen scholing heeft krijgt geen baan en wie het wel heeft ook niet.

Deze voorbeelden bewijzen niets, maar illustreren dat:

a) de kwaliteit van onderwijssystemen onder druk staan mede als gevolg van bezuinigingsmaatregelen;
b) de Amerikaanse high school geen perfecte oplossing biedt voor het vraagstuk van kwaliteit en effectiviteit van het onderwijs;
c) de aansluiting tussen onderwijs en arbeidsmarkt evenzeer problematisch is.

Tenslotte lijkt de man of vrouw in de straat het geloof in de lokale wetenschap en politiek te verliezen of verloren te hebben.

Kan de onderwijsresearch helpen mede antwoorden op deze problemen te vinden? In Boston dit keer geen centraal thema, waaronder alle bijdragen, na een selectie van 50% van het aanbod, zouden moeten worden gebundeld. Eindelijk is men tot de juiste conclusie gekomen dat een jaarmarkt zonder titel ook recht van bestaan heeft. Dat valt ook af te lezen aan de zich steeds verder uitbreidende reeks Special Interest Groups (zogenaamde SIG's) die een grote variatie aan onderwerpen en thema's inbrengt, zonder overigens het werk van de grote divisies van de AERA aan te tasten. Het is aan te bevelen dat veel jonge onderzoekers gebruik gaan maken van deze SIG's, omdat ze goedkoop zijn (meestal 5 à 6 $ per jaar), veel aan netwerkvorming doen en stimulerend werken. Het aantal leden van een SIG dat jaarlijks betaald heeft voor 15 juni is bepalend voor het aantal sessies dat een SIG mag organiseren op de AERA.

Opmerkelijk is elk jaar weer de participatie van Nederlandse onderzoekers tijdens de AERA. Ze zijn heel nadrukkelijk aanwezig, als paperlezers, als discussianten, als voorzitters van sessies en als (kritische) waarnemers. Dat geldt ook voor assistenten in opleiding (AIO) die de kans krijgen eigen werk te presenteren en zo via een internationaal forum kritiek ontvangen op hun werk. Een echte vorm van internationalisering, mede geconditioneerd door subsidies vanwege NWO en het van de USA geleende publish or perish principe: de beoordeling van VFO-programma's dwingt iedereen tot het bijeen sprokkelen van punten. Hopelijk heeft de wetenschap er ook iets aan.

In deze kroniek weer een min of meer toevallige samenstelling van impressies, waarbij getracht wordt zoveel als mogelijk hoofdindrukken van belangrijke divisies aan bod te laten komen. De selectie blijft echter persoonlijk en de oordelen zijn voor rekening van de auteurs.

Biesta en Miedema gaan in op recente ontwikkelingen in het Amerikaanse pedagogendebat. Mulder schetst de dilemma's van een nationaal curriculum, terwijl Van den Berg ingaat op de problemen rondom het schoolboek. De Klerk doet verslag van zijn ervarin-

Pedagogische Studiën 1990 (67) 450-469

gen m.b.t. divisie C: leren en instructie. Er lijkt verzadiging op te treden, al lijkt er ook perspectief te zijn voor een novice-novice paradigma. Wubbels toont ook enige scepsis wanneer hij het déjà vu laat klinken. Over de goede leerkracht is al heel lang geleden veel gezegd. Toch ook enkele nieuwe ontwikkelingen. Dan de effectieve school, door velen kennelijk niet goed begrepen. Van der Grift schetst enkele vooroordelen en misvattingen. Hij pleit voor een betere theorie over effectieve scholen. Nijhof gaat in op enkele ontwikkelingen in het Amerikaanse beroepsonderwijs die, jammer genoeg voor Europa, weinig nieuws bevatten. De beschrijving van enkele effectstudies op het gebied van de bedrijfsopleidingen roept bij hem de vraag op naar betere designs en effectmaten. Van Hout vervolgt zijn bijdrage van vorig jaar met een schets van een drietal grote problemen in het Amerikaanse Hoger Onderwijs en trekt een aantal interessante conclusies voor onze situatie. Ten slotte schetsen enkele CITO-medewerkers hun meest recente ervaringen tijdens de AERA en aanverwante organisaties.

1990 lijkt een reflectief jaar te zijn, waarin – naar het zich laat aanzien – niet echt nieuwe zaken zich doorzetten en het oude niet afgeschreven is. April 1991 zal bewijzen of een nieuwe lente ook een nieuw lied zal brengen in Chicago.

Ontwikkelingen in het Amerikaanse theoretisch-pedagogische debat (G. J. J. Biesta, Rijksuniversiteit Groningen en S. Miedema, Rijksuniversiteit Leiden)

In het navolgende geven we een indruk van de ontwikkelingen in het Amerikaanse theoretisch-pedagogische debat en laten we zien tot welke specifieke vragen voor historisch onderzoek dit leidt.

De meest opvallende ontwikkeling is naar onze mening de definitieve doorbraak van de *critical pedagogy*. De namen van Freire, Apple, Giroux en McLaren waren op de meest uiteenlopende plaatsen te horen, zowel bij theoretici als bij practici. Ofschoon de waardering voor hun werk sterk uiteenloopt, is het evident dat men niet meer om hun ideeën heen kan. Reynolds en Pinar merkten in hun bijdrage over de recente geschiedenis van de curriculum-theorie op dat de 'critical pedagogy' (ook bekend als 'new sociology of curriculum' en 'radical curriculum theory') heeft geresulteerd in "an almost taken-for-granted view that curriculum can be understood in any comprehensive sense only if it is contextualized socially, economically, and politically". Een opmerkelijke prestatie, omdat men in de literatuur tot aan het begin van de jaren zeventig nog algemeen uitging van de politieke neutraliteit van het curriculum. Nadat de critical pedagogy zich in de jaren zeventig vooral manifesteerde als een marxistisch-sociologische theorie over de reproducerende werking van de economische en ideologische basis in het onderwijs, kwam in het begin van de jaren tachtig de notie 'verzet' centraal te staan, zij het aanvankelijk nog vaak gecombineerd met een reproduktie-theoretische opvatting. Apple heeft in een vroeg stadium al opgemerkt dat het niet volstaat om alleen maar onderzoek te doen naar verzet, maar dat er feitelijk verzet gepleegd dient te worden. Via die 'beweging' komt, aldus Reynolds en Pinar, de critical pedagogy aan het einde van de jaren tachtig uit bij de politieke én de pedagogische praxis.

Met de groeiende belangstelling voor de critical pedagogy is ook de kritiek erop sterker geworden. Reynolds en Pinar lieten zien dat de critical pedagogy van haar Marxistische uitgangspunten is 'weggeschoven' in de richting van een liberale (Strike), een democratisch-humanistische (Bowers) c.q. een progressief-liberale (Wexler) positie. Dat doet de vraag ontstaan of en zo ja op welke basis critical pedagogy nog kritisch kan zijn. Het zijn de zich post-structuralistisch noemende pedagogen die de wending naar de praxis zoals die in de critical pedagogy is voltrokken, op theoretisch vlak hebben geradicaliseerd. Naar hun mening is er niet alleen op praktisch maar ook op theoretisch niveau sprake van een niet tot één rationeel principe te herleiden pluraliteit. Hiermee wordt dus de mogelijkheid voor een vaste (Marxistische) basis voor kritiek onderuit gehaald. Daarnaast verwijten o.a. Wexler en de feministische pedagoge Elsworth de academische kritische pedagogiek de aansluiting met de sociale bewegingen buiten de universiteit te missen. Academisch professionalisme wordt ten onrechte beschouwd als waarlijk politiek activisme. Wexler en Elsworth willen mét de sociale bewegingen een start maken voor 'tegenpraktijken'.

De goeroe van de Amerikaanse post-structuralistische pedagogen is zonder twijfel Der-

rida. Zowel Whitson, Stanley als Cherryholmes probeerden in hun bijdragen een (eigen) invulling te geven aan Derrida's dictum: 'Er is niets buiten de tekst'. Volgens Whitson betekent dit dat taal en werkelijkheid onlosmakelijk met elkaar zijn verweven, zodat er geen 'buiten' is. Deconstructie is daarmee ook geen negatieve, maar een affirmatieve act (Nietzsche). Hij zag het *poststructuralisme* voornamelijk als een aanvulling op de 'ideologiekritiek' van de Frankfurter Schule. Het post-structuralistische perspectief laat zien dat ideologie (ook) middels talige en grammaticale structuren tot stand komt en in stand blijft. Stanley begreep 'Er is niets buiten de tekst' als 'Er is niets buiten de *con*text'. Er is geen absoluut punt van waaruit kritiek geleverd kan worden, maar de context wordt wel bevraagd en gecritiseerd. Daar Derrida geen a priori sociaal-politiek programma heeft, is deze 'undecidability' een noodzakelijke voorwaarde voor het bepalen van de sociaal-politieke betekenis van gebeurtenissen en situaties. Kritiek kan slechts vanuit de ene 'context' op de andere worden geleverd. Met Derrida kom je dus *niet* uit bij een 'relativistische euforie', zo concludeerde Stanley. Derrida's project is een 'quest for better interpretations', een project waarbinnen de notie 'emancipatie' overigens haar betekenis kan behouden. Volgens Stanley impliceert emancipatie dat het onbetwijfelde in twijfel getrokken kan worden. Precies dát kan deconstructie doen. Deconstructie veronderstelt daarmee de zin en de mogelijkheid van kritiek, zij het dat die mogelijkheid altijd gebonden is aan een bepaalde context. Die context zelf bestaat uiteraard niet zomaar, maar is, zo besloot Stanley, altijd gebonden aan een bepaalde gemeenschap.

Cherryholmes plaatste net als Stanley de post-structuralistische aanpak binnen het project van de Verlichting. In tegenstelling tot 'klassieke' Verlichtingsprojecten als marxisme en feminisme is deconstructie naar zijn mening een niet-oppositionele werkwijze. Er wordt namelijk niet op voorhand een bepaald gelijk geclaimd. De post-structuralist dient de 'main-stream texts' van de pedagogiek én de opvoedingspraktijk te deconstrueren. Op die manier kan hij wijzen op ambiguïteiten, contradicties en paradoxen.

Sultan had vooral bedenkingen tegen de neutraliteit van de deconstructieve methode. Iedereen zou deconstructie voor zijn eigen doelen kunnen gebruiken. De methode is politiek neutraal en daarom, zo stelde Sultan, gevaarlijk. Het vereist een 'commitment to solidarity' wil deconstructie inderdaad tot radicale veranderingen kunnen bijdragen. Mogelijkheden voor deconstructie zag hij eerder langs Habermasiaanse lijnen. De taal- en machtsanalyses kunnen dienen voor handelingscoördinatie, werkelijkheidsontsluiting en het analyseren van machtsverhoudingen in sociale relaties. Dit werd echter niet geconcretiseerd.

Precies hier lag de betekenis van onze eigen bijdragen. Onze poging om de contouren van een *kritisch-pragmatische pedagogiek* te schetsen, was bedoeld om te laten zien hoe een waardevolle kritische positie (Habermas) *inhoudelijk* aangevuld zou kunnen worden (Rorty, Dewey) en aan kritische kracht zou kunnen winnen door de inbedding in de praxis duidelijk te maken én te realiseren (Dewey, Bernstein). Cherryholmes betwijfelde hierbij begrijpelijkerwijs of de 'redding' van een nieuw paradigma zou moeten komen. Ook McLaren benadrukte dit. Het 'systeemdenken' van Habermas (en ook Rorty) biedt naar zijn mening weinig perspectief. Juist 'the essential contingency of the universal' dient geaccentueerd te worden. Net als wij, maar ook net als bijvoorbeeld Apple en Stanley (en minder expliciet Cherryholmes), stelde McLaren uiteindelijk echter wel de praktische pedagogische vraag: "How to build a community?". Wanneer kritiek contextueel is, en de context gebonden is aan de gemeenschap, komt men uit bij de vraag hoe in de praktijk een zo ideaal mogelijke gemeenschap gerealiseerd en in stand gehouden kan worden. Een punt waar uiteraard ook de kritiek van Wexler en Elsworth relevant wordt.

Het bovenstaande maakt duidelijk dat vanuit verschillende posities uiteindelijk thema's worden opgeworpen (het verwerpen van vaste fundamenten voor kritiek) en vragen worden gesteld ("How to build a community?") die tot het erfgoed van het pragmatisme, met name dat van Dewey, behoren. Zowel in de hedendaagse critical pedagogy als vanuit post-structuralistisch perspectief als vanuit de continentale kritische theorie, wordt teruggegrepen op Deweys werk. Steeds wordt daarbij geprobeerd het kritisch gehalte van Deweys pragmatisme te verhogen (ofschoon iemand uit het publiek zich – wellicht terecht – afvroeg of 'critical pragmatism' eigenlijk geen pleonasme

was). Zo een synthese-poging is echter niet nieuw. Stanley betoogde dat de vertegenwoordigers van de 'social reconstructionists movement' (o.a. John Childs, George Counts, Theodore Brameld, Harold Rugg en Sidney Hook) reeds vanaf de twintiger jaren hebben geprobeerd om de pragmatistische pedagogiek meer kritisch te maken. Stanley liet zien hoe Brameld, in een poging om pragmatisme en marxisme aan elkaar te verbinden, dezelfde problemen aan de orde stelde en in de zelfde richting dacht als de vertegenwoordigers van de critical pedagogy nu. Hij meende dat met de ideeën van Brameld en de andere social reconstructionists een waardevolle bijdrage geleverd zou kunnen worden aan de hedendaagse discussies. Om dat te kunnen doen is volgens Stanley echter gedetailleerd historisch onderzoek nodig naar deze toch wat 'verloren' generatie van de Amerikaanse pedagogiek.

Terwijl Stanley om positieve redenen teruggreep op de eigen Amerikaanse traditie gingen er ook stemmen op om het continentale kritische gedachtengoed radicaal af te wijzen zodat men ex negativo bij de eigen kritische traditie uitkomt. Waks verdedigde de stelling dat William James en Theodore Brameld *wel* contact met 'het publiek' hadden en Henry Giroux *niet*, omdat deze laatste gebruik zou maken van 'esoterische Europese bronnen' (zoals Foucault, Derrida en Habermas) die als filosofische reconstructie van de Europese ervaring onmogelijk van betekenis zouden kunnen zijn voor Amerika. Waks' filosofisch-pedagogisch etnocentrisme kreeg echter niet veel bijval.

Het bovenstaande laat zien dat de Amerikaanse pedagogiek in een fase is beland waarin men de eigen kritisch-pedagogische historie gaat (her)ontdekken. Of de Amerikaanse kritische pedagogiek ook in Nederland een rol van betekenis zal gaan spelen, zodat het kritische denken, na aanvankelijk alleen uit het oosten nu ook vanuit het westen tot ons zal komen, zal in de toekomst moeten blijken.

Een nationaal curriculum? (M. Mulder, Universiteit Twente)

In de jaren tachtig is de discussie over nationale curricula versterkt voortgezet. In de Verenigde Staten bestaat geen curriculum op federaal niveau, ofschoon er een duidelijke trend is in deze richting. President George Bush heeft een reeks algemene nationale onderwijsdoelen voor het jaar 2000 vastgesteld. Uit de Gallup-opiniepeilingen van 1989 blijkt dat 70% van de Amerikaanse bevolking voorstander is van nationale eindtermen. Eveneens 70% is voor een nationaal curriculum en 77% voor een systeem van landelijke examinering. Op dit moment zijn er peilingsonderzoeken in 37 staten, er is een modelcurriculum voor het primair en secundair onderwijs vervaardigd en er worden landelijke examens voor de lerarenopleiding ontwikkeld. Smith constateerde dat er nationale doelen, inhouden, raamleerplannen, toetsen, eindtermen en inhouden voor de lerarenopleiding zijn en vroeg zich af hoe lang het nationale curriculum nog op zich kan laten wachten. Hij gaf twee argumenten voor een nationaal curriculum: de verantwoordingsplicht enerzijds en de vergelijkbaarheid van scholen anderzijds. Als scholen zich moeten verantwoorden voor hun resultaten, dan moeten zij eerst consensus bereikt hebben over het curriculum. En als de prestaties en pedagogisch-didactische werkwijze fair moeten kunnen worden beoordeeld, dient tevoren een gemeenschappelijk curriculum aanwezig te zijn, anders ontbreekt een vergelijkingscriterium. Een vraag voor de jaren negentig is op welke wijze de top-down benadering bij de ontwikkeling van een nationaal curriculum kan worden verbonden met lokaal initiatief en de inspanning om onderwijsgevenden te motiveren om dat curriculum uit te voeren.

Volgens Shanker dient het onderwijs in de Verenigde Staten het vertrouwen van het publiek te herwinnen door gehoor te geven aan de druk van ouders in de richting van een ander systeem van accountability. Daarvoor is in ieder geval nodig dat er een nationaal curriculum wordt vastgesteld. Het belangrijkste argument dat hij hiervoor naar voren bracht was de mobiliteit van de bevolking. Door verhuizingen van leerlingen ontstaan er ten gevolge van het ontbreken van een nationaal curriculum fricties in het leerproces. Een nationaal curriculum is noodzakelijk, maar hoever moet deze nationalisatie gaan? In Californië kent men momenteel een leerplan dat fungeert als guideline, landelijke examinering en periodieke peiling van de onderwijskwaliteit, een systeem dus dat op een aantal punten overeenkomt met het Nederlandse stelsel voor funderend onderwijs.

Geiger daarentegen typeert het nationaal curriculum als een oxymoron: een landelijk vastgelegde lessentabel en lokale autonomie zijn begrippen die diametraal tegenover elkaar staan. Dominante begrippen in de onderwijsvernieuwing van de afgelopen jaren waren 'local control', 'site-based decision making', 'teacher empowerment' en 'school restructuring'. In 1986 zijn drie zeer invloedrijke documenten gepubliceerd die indicatief waren voor de onderwijsvernieuwing in die jaren: De National Governers' Association publiceerde "Time for Results", de Carnegie Forum on Education and the Economy publiceerde "A Nation Prepared", en de Education Commission of the State publiceerde "What Next? More Leverage for Teachers". De rapporten speelden een belangrijke rol in het formuleren van de volgende conclusies:

- top-down vernieuwing is een recept voor mislukking;
- schoolmanagement moet worden gedecentraliseerd en het onderwijs moet intern worden gedemocratiseerd;
- plaatselijke initiatieven en lokale autonomie zijn bepalend voor effectieve vernieuwingen.

Sinds de publikatie in 1983 van "A Nation at Risk" had het onderwijs steeds het gevoel dat voorstellen voor onderwijsvernieuwing te ver af stonden van de praktijk, terwijl de publikatie van genoemde rapporten een gevoel van herkenning oproept. Het onderwijs had reeds bij voortduring gesteld dat hervormingen volgens een bottom-up-benadering zouden moeten plaats hebben, dat democratische besluitvormingsstructuren nodig waren en dat de invloed van onderwijsgevenden en andere direct bij het onderwijs betrokkenen op de besluitvorming zou moeten worden versterkt. Volgens deze spreker is er wel behoefte aan nationale onderwijsdoelen, maar de lokale autonomie van het onderwijs moet daarbij worden gerespecteerd.

Ook Cross betuigde zich tegenstander van een nationaal curriculum. Hij stelde dat er weliswaar een trend zichtbaar is in de richting van een nationaal curriculum: de verzamelde leerboeken, die in twintig staten door leerboekselectiecommissies per staat worden goedgekeurd, fungeren als een soort informeel curriculum, terwijl ook de periodieke peilingsonderzoeken een sturende werking hebben in de richting van een nationaal curriculum. Een nieuw soort periodiek peilingsonderzoek wordt echter ontwikkeld waardoor scholen worden geholpen het onderwijs sterker te richten op haar essentiële doelen waaronder wiskundig leren denken, leren probleemoplossen en het leren toepassen van historische en geografische kennis. Beslissers op alle niveaus krijgen daarbij dan betere feedback over de leerprestaties. Als dit systeem van periodiek peilingsonderzoek is gerealiseerd, is er volgens de spreker geen behoefte meer aan een nationaal curriculum. Het staat bovendien op gespannen voet met de tendens van verdergaande decentralisatie in de Verenigde Staten. Op centraal niveau bestaat uitsluitend belangstelling voor de uitkomsten van het onderwijs en niet voor het curriculum als zodanig. Dat maakt namelijk deel uit van het onderwijs en is niet het produkt van het onderwijssysteem. Er kunnen dus twee verschillen doch complementaire trends worden gesignaleerd: decentralisatie van het onderwijsproces en alles wat daarmee samenhangt enerzijds, en centralisatie van de informatie over het resultaat van het onderwijs.

Lokale autonomie en zelfbeheer

King rapporteerde over een studie naar de wijze waarop in de staat Minnesota wordt omgegaan met de spanning tussen lokale autonomie en centraal beheer. Genoemde staat leent zich goed voor de beschrijving van dit probleem omdat het onderwijs aldaar wordt getypeerd door haar sterk innovatieve karakter en decentralisatie enerzijds en centraliserende tendensen anderzijds. Recente centralistische wetgeving is gericht op het vergroten van de schoolkeuze van ouders en leerlingen en de zogenaamde outcome-based education. Door het inbouwen van marktmechanismes denkt de overheid concurrentie en daarmee kwaliteitsverhoging te bevorderen, ofschoon de grote vraag is of ouders over voldoende informatie beschikken om op verantwoorde wijze een school te kiezen voor hun kinderen. Uit onderzoek is gebleken dat circa 40% van de ouders uit puur praktische overwegingen een school kiest. Het is daarom niet zeker of genoemde regelgeving zonder meer leidt tot de beoogde resultaten. Ook bij de outcome-based education kan dezelfde twijfel worden geuit. De staat heeft wel het beheer over de geformuleerde outcomes, maar de scholen staan vervolgens voor de opgave deze outcomes om te zetten in nieuwe curricula en instructieprocessen. Voor

velen lijkt deze benadering te veel op de golf van meetbare leerdoelen, waarvan is gebleken dat het tot weinig of geen verbetering van het onderwijsresultaat heeft geleid. Van een ander initiatief, namelijk de oprichting van een praktisch gericht onderzoeksinstituut, het CAREI, wordt meer verwacht. Het doel van dit instituut, dat in het leven is geroepen door een aantal districten in Minnesota, is actie-onderzoek te doen in samenwerking met de scholen in de desbetreffende districten, om reflectie op de bestaande praktijk te bevorderen en heel gericht situatie-specifieke praktische adviezen te genereren voor de transformatie van het onderwijs. Dit initiatief is een teken dat verschillende districten willen uitstijgen boven de regelgeving op staat-niveau. Als potentieel probleem werd gesignaleerd dat de ontmoeting van de onderzoeks- en de praktijkcultuur erg lastig kan zijn vanwege de verschillende belangen van onderzoekers en praktici.

Silberstein en Sabar gingen eveneens in op het spanningsveld tussen centraal beheer en lokale autonomie. Zij rapporteerden over een studie in Israël naar de zogenaamde autonome scholen die op creatieve wijze centrale verplichtingen en lokale rechten wisten te combineren, waarbij zij geen ogenschijnlijk dwingende keuze maakten voor een bepaalde onderwijskundige ideologie. Substantieel verschillende uitgangspunten werden vertaald in curriculumcomponenten en consistent uitgewerkt in de onderwijspraktijk. Hierdoor konden de desbetreffende scholen twee klassieke valkuilen vermijden: het 'pendulum' fenomeen, de vicieuze cirkel van centralisatie en decentralisatie als een voortdurend actie- en reactieproces, en het 'kloof'-fenomeen, de frictie die bestaat tussen het formeel en het operationeel curriculum.

Angus schetste een drietal problemen die voortvloeiden uit het decentralisatiebeleid in het onderwijs in West-Australië. In de eerste plaats bleek het moeilijk te zijn te bepalen waar de grenzen liggen van de autonome school in een gedecentraliseerd onderwijsbestel. Ten tweede percipieerde het onderwijsveld de decentralisatie ook als bedreiging van de lokale autonomie, omdat decentralisatie gepaard gaat met centrale regelgeving. Men vreesde daarom dat door het recente beleid het onderwijs zou verworden tot een speelbal van landelijke politieke besluitvorming. Ten derde trad er een onbedoeld effect op. Door het decentralisatiebeleid en de daarmee gepaard gaande herstructurering van de financiering van het onderwijs werd het beeld versterkt dat schoolverbetering wordt bevorderd door het toevoegen van functies aan het bestaande onderwijs en het veranderen van structuren, in tegenstelling tot het herontwerpen van het werk dat scholen verrichten.

Zelfevaluatie

Bjorndal ging in op de consequenties van het delegeren van bevoegdheden in het zogenaamde 'upper-secondary' onderwijs in Noorwegen en beschreef een project over praktijkgerichte zelfevaluatie van scholen. Om binnen het kader van het decentralisatieproces de regionale en lokale autonomie in overeenstemming te houden met de nationale doelen voor het onderwijs en een landelijk gelijk niveau van onderwijskwaliteit te waarborgen, is veel aandacht besteed aan in-service opleiding van het management en samenwerking binnen scholen. Voorts is een project uitgevoerd ten behoeve van organisatie-ontwikkeling van onderwijsinstellingen. Parallel hieraan is een project gestart voor schoolevaluatie. Hierin moest een evenwicht worden gevonden tussen de nationale belangen en de lokale autonomie. Er is vastgesteld dat de evaluatie gekenmerkt diende te worden door haar systematiek en lange termijn gerichtheid. De individuele school diende het middelpunt te zijn van de evaluatie, terwijl de nationale doelen als evaluatie-context dienen. Bij de start van het project werd de soms te geringe evaluatie-expertise als probleem ervaren. Andere problemen die werden genoemd, zijn: de kennis over schoolevaluatie is nog beperkt evenals de voor de zelfevaluaties benodigde tijd, het gevaar dat het evaluatieproces als ontwikkelingsproces wordt opgevat waardoor er feitelijk geen evaluatie meer plaats heeft, de vrees dat in de verplichte jaarverslagen gevoelige resultaten worden opgenomen die men liever vertrouwelijk wil behandelen en de vrees voor centrale kwaliteitsmetingen die een bedreiging kunnen vormen voor de lokale autonomie.

Het schoolboek, criteria en de markt (E. van den Berg, Universiteit Twente)

In onze westerse samenlevingen zijn scholen zonder schoolboeken nauwelijks voor te stel-

len. Misschien is juist deze vanzelfsprekendheid er de oorzaak van dat binnen het onderwijskundig onderzoek vrij weinig belangstelling is voor schoolboeken. Met de uitgave van het 89e jaarboek van The National Society for the Study of Education "Textbooks and Schooling in the United States" (Eds.: D. L. Elliott & A. Woodwand; uitgever: University of Chicago Press) is een poging ondernomen om het 'schoolboek' weer wat hoger op de agenda van het onderwijsonderzoek te plaatsen. In Boston vond de presentatie van dit boek plaats. De Special Interest Group 'Textbooks and Textbookpublishing' wijdde haar jaarlijkse vergadering geheel aan de uitgave van dit boek.

Na een aantal korte inleidingen en lovende woorden volgde een pittige discussie over de rol en invloed van educatieve uitgevers op het schoolsysteem. Telkens speelden onderwijsonderzoekers aan de ene kant en educatieve uitgevers aan de andere kant elkaar de zwarte piet toe.

Leesbaarheidsformules

Eén van de discussiepunten was het hanteren van leesbaarheidsformules. Jeanne Chall had in haar inleiding scherpe kritiek geuit op het rigide gebruik van deze formules door educatieve uitgevers. Refererend aan onderzoek stelde zij dat vanaf de late jaren twintig tot heden de taal in schoolboeken steeds eenvoudiger is geworden. Volgens haar was het in de jaren dertig noodzakelijk om de taal in schoolboeken te vereenvoudigen o.a. vanwege de komst van immigrantenkinderen op de Amerikaanse scholen voor wie het Engels niet de moedertaal was. Zij is echter van mening dat het vereenvoudigen van de taal in schoolboeken te ver is doorgeschoten. Zij spreekt over 'dumbing of textbooks'. De educatieve uitgevers stelden in hun verweer dat niet zij voor dit proces verantwoordelijk zijn, maar de selectiecommissies voor schoolboeken van staten en districten. Zij plaatsen immers schoolboeken met een te moeilijk taalgebruik niet op de boekenlijst waaruit de scholen kunnen kiezen. Dus wat moeten zij als uitgevers anders dan aan de wensen van de markt voldoen.

Richtlijnen

Een ander discussiepunt was naar aanleiding van de plannen van de staat Californië om stringente richtlijnen op te stellen waaraan de schoolboeken voor deze staat moeten voldoen. De discussie spitste zich toe op de vraag of het wenselijk is om dergelijke richtlijnen per staat of zelfs op nationaal niveau vast te leggen. De onderwijsonderzoekers waren verdeeld. Sommigen zagen wel wat in het idee van door overheidswege vastgestelde criteria waaraan schoolboeken moeten voldoen. Hierdoor zou de macht van de educatieve uitgevers om feitelijk door middel van de schoolboeken het curriculum te definiëren doorbroken kunnen worden. Dit is volgens hen noodzakelijk omdat educatieve uitgevers vrij conservatieve produkten maken die een handicap vormen wanneer men onderwijsvernieuwingen wil realiseren. De uitgevers waren het hier uiteraard niet mee eens. Zij stelden dat hun produkten voldoen aan de eisen die de markt stelt. Als er voor een onderwijsvernieuwing een voldoende grote markt is dan leveren zij de materialen conform de wensen van de markt. Als er kritiek is op de onderwijskundige kwaliteit van schoolboeken dan kan dit slechts betekenen dat de critici te laag in aantal zijn om een voldoende grote markt te vormen of wanneer het aantal wel voldoende groot is, zij hun wensen niet duidelijk genoeg kenbaar maken. Andere onderzoekers zagen niet veel in het afleiden van criteria waaraan schoolboeken moeten voldoen uit nationale curricula. Zij vreesden dat dit zou leiden tot uiterst formele criteria die eenzelfde uitwerking kunnen krijgen als het rigide toepassen van leesbaarheidsformules.

Ook de uitgevers waren geen voorstander van richtlijnen waaraan hun produkten moeten voldoen. Zij vreesden eenvormigheid of zoals één van hen het formuleerde: 'Dan valt alleen nog te concureren met nog mooiere plaatjes.' Maar als er toch richtlijnen moeten komen dan op nationaal niveau. De uitgevers menen dat de produktie van schoolboeken voor afzonderlijke staten uit financieel oogpunt niet haalbaar is. De plannen van de staat Californië werden dan ook beslist niet met gejuich ontvangen. Een vergelijking met de Nederlandse situatie dringt zich op. Californië, met ruim 28 miljoen inwoners, wordt door de grote Amerikaanse educatieve uitgeverijen als een te kleine markt beschouwd. Nederland telt nog niet de helft van dit aantal inwoners, terwijl er ongeveer 30 uitgeverijen lid zijn van de Groep Educatieve Uitgeverijen.

Hoewel de situatie in de Verenigde Staten niet helemaal te vergelijken is met de Nederlandse,

is de discussie over de rol van educatieve uitgeverijen in het onderwijssysteem ook hier actueel. Zeker omdat in Nederland vanwege de grondwettelijke vrijheid van richting en inrichting van onderwijs de overheid geen enkele bemoeienis mag hebben met de keuze van schoolboeken door de scholen. Ook in Nederland geldt dat er een spanning bestaat tussen wat onderwijsvernieuwers graag in schoolboeken willen zien en hoe uitgevers schoolboeken inhoud en vorm geven. Deze laatste groep laat zich leiden door de markt, die niet zo vernieuwingsgezind is als velen graag zouden willen.

Leren en instructie (L. F. W. de Klerk, Katholieke Universiteit Brabant)

Leerling- versus leraar-gestuurde instructie
Een interessant poster-symposium werd verzorgd door Palincsar over het onderwerp: leerling- versus leraar-gestuurde instructie. Hierbij werden drie instructie-procedures met elkaar vergeleken: 'directed learning' (waarbij het leerproces gestuurd wordt op basis van activiteiten van de leerling); 'reciprocal learning' (waarbij de leraar voornamelijk als informant optreedt); en 'collaborative learning' (waarbij de leerlingen initiatieven nemen ten aanzien van de te volgen werkwijze). Het accent lag in dit onderzoek niet zozeer op de vraag welke conditie de beste resultaten levert, maar veeleer op de vraag wat er precies gebeurt binnen iedere procedure in termen van regulatieprocessen. Hiertoe werden verschillende groepjes kinderen gedurende lange tijd intensief geobserveerd. Ook werden behalve prestaties allerlei leerlingkenmerken gemeten, zoals voorkennis, strategiegebruik en metacognitieve vaardigheden. Gebleken is dat er complexe interacties optraden tussen leerlingkenmerken, instructieprocedure en resultaten. Een belangrijke conclusie die uit dit onderzoek naar voren kwam is dat verschillen tussen procedures enigszins gecamoufleerd worden doordat leerlingen ook hun eigen doelen en werkwijzen kennen, die soms interfereren met die van de instructies. Ook is gewezen op het belang van de sociale context waarin het leren plaats vindt. Deze conclusie werd onderschreven door Resnick. Volgens haar wordt kennis niet louter 'geconstrueerd' (vgl. constructivisme) door middel van instructie, maar hangt het resultaat van instructie mede af van hoe de leerlingen de leertaak interpreteren en waarderen. Zij merkt op dat 'human life is a passage through activity situations', waarmee zij verwijst naar de 'situated cognition'-visie op leren en instructie. Volgens Resnick is het zo dat de ervaringen met situaties bepalen hoe de leerling een omgeving interpreteert, hetgeen weer bepalend is voor zijn motivatie. Zijn prestatieniveau hangt echter niet alleen af van de motivatie, maar ook van de mate waarin hij in staat is gebleken de genoemde ervaringen te 'integreren'.

Situated cognition and learning paradigma
'Situated cognition' is in verschillende sessies ter sprake gekomen. Deze term is al weer een paar jaar in omloop. Ook in de kroniek naar aanleiding van de AERA- conferentie van 1989 is hieraan reeds aandacht besteed. De aanhangers van deze benadering zien hierin een herleving van het bekende informatieverwerkings-paradigma. Kenmerkend voor de traditionele informatie-verwerkingsopvatting over leren en instructie is de vraag via welke cognitieve operaties informatie verwerkt wordt. Door middel van onderzoek wordt nagegaan hoe leerlingen informatie opnemen, verwerken, opslaan in het geheugen en het daaruit terughalen (retrieval) voor gebruik, bijvoorbeeld om er bepaalde problemen mee op te lossen. Kennis van de werking van het cognitieve systeem levert de grondslag voor het ontwerpen van effectieve leersituaties. Bij het 'situated cognition and learning' paradigma ligt het accent niet op het cognitieve systeem, maar op de situatie waarin geleerd wordt. Het gaat hierbij niet om wat een leerling weet, maar om wat hij kan in een gegeven situatie. Deze opvatting heeft geleid tot een andere visie op leren en instructie, een visie die aangeduid wordt met de term 'cognitive apprenticeship'. Dit model is erop gericht "to enculturate students into authentic processes through activity and social interaction in a way similar to that evident – and evidently successful – in craft apprènticeship" (om de woorden van J. S. Brown te citeren, die de term situated cognition gelanceerd heeft). Volgens dit type onderwijs-leersituatie moet ernaar gestreefd worden een adequate leeromgeving te creëren, waarin veel hulpbronnen, 'tools', leermaterialen en sociale interactiemogelijkheden aanwezig zijn, op een wijze die representatief is voor toekomstige werk- of

taak-situatie.

Niet alle onderzoekers zijn even hoopvol gestemd over de mogelijkheden die 'situated cognition' te bieden heeft. Gavriel Salomon toonde zich een krachtig opponent. Hij is van mening dat deze term een modewoord is, dat tot situaties leidt waarbij je maar moet afwachten wat er gebeurt. Er zal veel onduidelijkheid blijven bestaan over wat, hoe, waarom en hoe goed geleerd wordt. Door grote nadruk te leggen op de sociale interactie worden (te) hoge eisen gesteld aan de didactische kwaliteit van de leraar (expert). Hij moet een model zijn en tegelijkertijd rekening houden met individuele verschillen. Als het aantal leerlingen groot is dan dient hij bovendien over de juiste management-kwaliteiten te beschikken. Echter, een goede 'leermeester' zal de neiging hebben zijn aandacht vooral te richten op die leerlingen die dat nodig hebben. Dit zal de effectiviteit en de efficiëntie van de leersituatie – criteria die centraal staan in het onderzoek dat is opgezet volgens het traditionele informatieverwerkingsparadigma – niet ten goede komen. Volgens Salomon gaat het bij 'situated cognition' om veel geschreeuw en weinig wol!

Novice-novice paradigma

Opmerkelijk is dat zeer veel presentaties op het gebied van leren en instructie betrekking hebben op kleinschalig onderzoek, soms met een duidelijke vraagstelling, maar vaak ook exploratief van aard. Dit maakt het moeilijk om de stand van zaken op het terrein van divisie C goed in beeld te krijgen. Vooral als de separate onderzoeken binnen een thema niet in samenhang worden gepresenteerd – hetgeen nogal eens voorkwam – dan is het moeilijk in te schatten wat de bijdrage is aan het betreffende kennisdomein. Ook is het vaak moeilijk te beoordelen of de effecten die in experimenten gevonden worden ook gelden voor het onderwijs en welke maatregelen daarin getroffen moeten worden om dezelfde effecten te bereiken.

Een voorbeeld van kleinschalig fundamenteel onderzoek waarbij wel gelet is op de onderlinge samenhang en op de betekenis die het onderzoek kan hebben voor zowel de research als de praktijk van het onderwijs is gepresenteerd en becommentarieerd tijdens een symposium over o.a. de rol van voorkennis bij het leren op school (Understanding children's memory development: the roles of knowledge, aptitude, and strategies). Eén van de onderzoekers die daarin een presentatie hield was Wolfgang Schneider van het Max Planck Instituut te München. De centrale vraagstelling die aan zijn onderzoek ten grondslag lag luidde: wat neemt een leerling waar indien hij geconfronteerd wordt met een bepaald type vraagstuk of probleem en waarvan is dat afhankelijk? Ziet een gevorderde leerling (of expert) meer of ook andere dingen dan een beginnende leerling (of novice)? Zoals De Groot dat in Nederland al veel eerder had gedaan heeft ook deze onderzoeker gekeken naar wat beginnende en gevorderde schakers zich herinneren van gepresenteerde stellingen. Uiteraard wisten de experts zich veel meer te herinneren dan beginners, maar dat gold voornamelijk voor 'realistische stellingen'. Zodra de stukken at random op het bord geplaatst worden, blijkt dit verschil aanzienlijk kleiner te zijn. Een dergelijk experiment is ook gedaan met betrekking tot voetbal (soccer). Dit onderzoek leverde gelijkluidende uitkomsten op: voorkennis bepaalt in hoge mate hoeveel en vooral ook wat onthouden wordt. Dit leidt tot de vraag wat men in het onderwijs moet doen. Moet men aansluiten bij kennisdomeinen waarop de leerling enigermate gevorderd is (moet men bij iemand die veel van voetballen weet aansluiten bij zijn kennis op dat terrein als het gaat om taal en rekenen; met andere woorden definieert voorkennis de context waarin het leren moet plaats vinden?) of moet men domeinspecifieke strategieën ontwikkelen bij de leerling (via oefening) waarmee hij in staat zal zijn de problemen uit het betreffende domein aan te pakken en op te lossen? Als niet alleen gekeken wordt naar verschillen tussen experts en beginners, maar ook naar verschillen tussen beginners onderling, dan blijkt niet alleen de door de voorkennis gedefinieerde context, maar vooral ook de sociale context (en in het bijzonder het samen werken met peers) van belang te zijn voor het leerproces. Een interessante conclusie is dat het expert-novice paradigma tot nog toe teveel het accent gelegd heeft op wat een expert weet en kan en dat van daaruit wordt afgeleid wat een novice tekort komt. Dat moet dan vervolgens door de instructie worden aangevuld. Maar, beginners zijn niet op alle terreinen onwetend of dom. Vergelijkt men de novice met een expert op een bepaald domein dan heeft men een eenzijdige belichting van de novice en bovendien komt men dan niet te

weten wat er buiten het betreffende domein te halen valt (of reeds aanwezig is) dat ook bijdraagt tot het verwerven van de beoogde competentie. Het invoeren van het reeds eerder in Europa door Elshout voorgestelde *novice-novice-paradigma* zou wel eens een interessante aanvulling kunnen zijn op de gebruikelijke wijze waarop taakanalyses worden uitgevoerd. Niet alleen letten op de einddoelen of op de 'prerequisite skills' die uit de einddoelen afgeleid kunnen worden, maar evenzeer op de beginsituatie en met name op individuele verschillen die er bestaan tussen beginners. Op welke algemene vaardigheden grijpen zij terug? Welk type strategieën gebruiken zij? Op welke specifieke ervaringen doen zij een beroep? Is er hierbij sprake van transfer? Als dit soort vragen beantwoord is dan doet zich vervolgens natuurlijk de vraag voor op welke wijze in het onderwijs met deze verschillen rekening gehouden moet worden. Of komt men zo via de sociale context weer terecht bij 'situated cognition'? Volgens de discussiant van het symposium niet. Hij kwam uit bij Emanuel Kant, althans voor zover het ging over de invloed van voorKENNIS op verder leren. En hiermee deed hij iets wat bij veel paper-presentaties te beluisteren viel: 'filosoferen'. Dit gebeurde met name door een aantal 'oude bekenden' (zoals bijvoorbeeld Greeno). In plaats van onderzoeksresultaten te presenteren werden allerlei beschouwingen gegeven over de zin van onderwijs, over de relatie onderwijs en samenleving, en over de pedagogische contexten waarin het onderwijs gesitueerd moet worden. In hoeverre hier sprake is van een verschijnsel dat het einde van een productieve onderzoeksperiode van de betreffende sprekers inluidt dan wel van een periode van bezinning die het begin van een nieuwe onderzoeksperiode aangeeft, is nog onduidelijk. De indruk bestaat helaas dat het eerste vaker het geval is dan het laatste.

Leraren en lerarenopleiding (Th. Wubbels, Rijksuniversiteit Utrecht)

In deze bijdrage bespreek ik mijn belangrijkste indrukken van de presentaties over leraren en lerarenopleiding. Eerst echter een chauvinistische opmerking. Voor de Nederlanders was het aardig om te constateren dat in de presentaties op dit gebied steeds vaker gesproken wordt over de fenomenologische aanpak. Deze uit Europa geïmporteerde benadering krijgt meer Noordamerikaanse volgelingen, met name dankzij het werk van Max van Maanen. Zelfs de door Bleeker en Mulderij ontwikkelde methoden worden gebruikt. Amerikaans provincialisme verhindert meestal nog dat er naar de oorspronkelijke Europese bronnen wordt verwezen. Het blijft bij het raadplegen van de Amerikaanse exegeses.

Constructivisme

De laatste jaren werd de term 'reflecteren' te pas en te onpas in sessies over leraren en lerarenopleiding gebruikt. Dit jaar was er een nieuwe kreet ('buzz-word') te horen: 'constructivisme'. Naar een constructivistisch theoretisch kader werd de laatste jaren al steeds vaker verwezen. Nu was er wat dit betreft een definitieve doorbraak en gaven veel presentatoren aan dat ze leraren opvatten als zelf, actief, kennis construerende subjecten, die daarbij bestaande cognitieve structuren gebruiken om nieuwe ervaringen te begrijpen. Aangezien de AERA-presentaties zeker niet vrij zijn van modegrillen is het te vroeg om te spreken van een nieuwe stroming in het onderzoek naar leraren en lerarenopleiding. Wellicht kunnen we over vijftien jaar vaststellen of het hier om een even substantiële beweging gaat als die rond reflecteren.

Het belijden van constructivistische uitgangspunten betekende geen eenheid in onderzoeksbenadering: sommigen gaven kwalitatieve, interpretatieve analyses van de wijze waarop leraren hun kennis construeerden, anderen rapporteerden vooral over processen in lerarenopleidingen of bij innovaties. Een specifieke groep presentaties hield zich op twee niveaus met constructivisme bezig: namelijk, met onderzoek naar constructivistisch onderwijs door leraren in de klas en met scholing om leraren te helpen les te geven volgens constructivistische uitgangspunten. Op dit tweede niveau spelen dezelfde uitgangspunten een rol als op het eerste. Hiermee zette een trend door, die in voorgaande jaren in presentaties van de Special Interest Group 'Subject Matter Knowledge and Conceptual Change' kon worden geconstateerd: men tracht het gebruik van strategieën voor het veranderen van leerlingendenkbeelden te verbreden naar het veld van de lerarenopleiding. Daarbij gaat het om vragen als: Met welke opvattingen (met name over goed onderwijzen) komen (aanstaande) leraren de opleiding

of nascholing binnen? In hoeverre zijn deze opvattingen in strijd met wat het kennisbestand over lesgeven zegt en/of met constructivistische uitgangspunten? En ten slotte: Op welke wijze kunnen de studenten geholpen worden concepten te construeren die meer met dat kennisbestand in overeenstemming zijn?

Enkele interessante bijdragen uit de school van Tobin (Tobin, Briscoe, Taylor) rapporteerden gevalsbeschrijvingen onder meer over het proces van overgang van een traditionele onderwijsvorm (de leraar die overdraagt) naar een meer constructivistische (de leraar die constructieprocessen bevordert). Er bleek weliswaar een meer leerlinggecentreerde opvatting te ontstaan, maar in de lesaanpak bleef de leraarsrol toch zo centraal dat de leerlingen onvoldoende actief eigen concepten konden verkennen.

O'Loughlin analyseerde eigen dagboeken over zijn pogingen constructivistisch onderwijs in de lerarenopleiding te geven. Hoewel zijn analyses interessant waren doet zich, net als bij veel verwant werk, de vraag voor wat zo'n analyse meer oplevert dan interessante gedachten van een toevallige opleider/onderzoeker. Dezelfde auteur presenteerde in een andere sessie zijn zienswijze op de subjectieve theorieën van leraren en de effecten daarvan op hun onderwijs. Hij trachtte via zijn analyse te laten zien dat leraren vaak meer mogelijkheden hebben om hun omgeving te veranderen dan ze denken. De beperkingen die ze zien zijn immers ook (slechts) constructies van de werkelijkheid. Er kunnen daarvoor andere constructies in de plaats komen die de leraar meer kans geven om tot veranderingen te komen. O'Loughlin legde een verband met het kritisch reflecteren, zoals dat in het opleidingsprogramma van Zeichner en Liston wordt nagestreefd. Hoewel dit opleidingsprogramma volgens O'Loughlin een constructivistisch uitgangspunt heeft en het beste in zijn soort is, acht hij het, net als Zeichner en Liston zelf, niet succesvol.

Tobin besteedde ook in zijn lezing ter gelegenheid van het winnen van de Raymond Catell Award veel aandacht aan het (radicaal) constructivisme. Een centrale stelling uit zijn betoog betrof het slagen van onderwijsveranderingen in de les. Op grond van werk in zijn onderzoeksgroep achtte hij een noodzakelijke voorwaarde daarvoor dat leraren een, voor hen persoonlijk, betekenisvol beeld opbouwen over hoe het curriculum er in de eigen lessen uit zou kunnen zien.

De bijdragen waarin het constructivisme centraal stond overziend, riep het geheel bij mij de gedachte op: oude wijn in nieuwe zakken. Veelal leiden de onderzoeken tot de conclusie dat een onderzoekende aanpak in het onderwijs of de lerarenopleiding nodig is. Dat leek verdacht veel op wat vele jaren geleden al werd voorgestaan door bijvoorbeeld Bruner. En net als toen is er daarbij onvoldoende aandacht voor de vraag hoe leraren het ontstaan van cognities (constructies) kunnen bevorderen. Verder moet geconstateerd worden dat veel onderzoekers nogal gemakkelijk generaliseerden op basis van werk dat exploratief genoemd moet worden; het ging veelal slechts om een beperkt aantal gevalsbeschrijvingen.

Cognities

In de kroniek van het vorige jaar constateerde Verloop dat impliciete theorieën steeds meer aandacht krijgen in de niet aflatende stroom van onderzoek naar cognities van leraren. Het lijkt er nu op dat studies waarin de leraar als probleemoplosser of beslisser wordt opgevat vrijwel geheel plaats hebben gemaakt voor pogingen om meer persoonlijke beschrijvingen van cognities te verkrijgen. Vooral beschrijvingen met behulp van metaforen waren erg in trek. Het gaat daarbij onder meer om onderzoek naar de manier waarop leraren middels eigen metaforen (zoals kapitein van het schip, entertainer) hun rol begrijpen en hoe die metaforen hun lespraktijk beïnvloeden. Andere papers rapporteerden over studies waarin de onderzoekers zich de taak gesteld hadden het denken van de leraren te beschrijven met, door de onderzoekers bedachte, metaforen. Het onderscheid tussen de metaforen die voor de onderzoekers bruikbaar zijn om de werkelijkheid te beschrijven en de daadwerkelijke leraarscognities werd daarbij wel eens uit het oog verloren.

De vele onderzoeken over de opvattingen die (aanstaande) leraren hebben over goed lesgeven en goede leraren leverden niets nieuws op. Vrijwel al het gepresenteerde onderzoek over leraarscognities maakte gebruik van interviews en observaties bij een gering aantal respondenten. Sommige papers gaven duidelijke beschrijvingen van dataverzamelings- en analysetechnieken, maar het aantal papers waarin de methodische verant-

woording vrijwel geheel ontbrak was groot. Met name een paper van Stoddart en Roehler vormde een gunstige uitzondering. Zij gaven een analyse van de veranderingen in kennisstructuren van aanstaande leraren op basis van een kwantitatieve aanpak, uitgaande van boomdiagrammen die de cognitieve structuur weergaven.

Opleidingsprogramma's

Tekorten aan leraren, vooral in de exacte vakken, alsmede zorgen om de kwaliteit van nieuwe leraren leiden tot verschillende ontwikkelingen. Aan de ene kant inspireren ze tot allerlei omscholingsprogramma's en nieuwe manieren om leraren aan universiteiten op te leiden. Aan de andere kant ontstaat een roep van scholen en schooldistricten om hen de vrijheid te geven op hun eigen manier in hun behoefte aan leraren te voorzien. Ze wensen zelf de verantwoordelijkheid voor de kwaliteit van de leraren te krijgen en daarmee, de bevoegdheid om leraren op te leiden. Dit heeft er toe geleid dat sommige schooldistricten nu zogenaamde alternatieve opleidingsprogramma's organiseren voor mensen zonder enige didactische scholing. In die programma's vormt het leren door te doen in de lespraktijk vrijwel het enige onderdeel. Het diploma wordt door het schooldistrict uitgereikt. De lerarenopleidingen hebben zich hier uit de markt geprijsd, doordat ze er niet in geslaagd zijn anderen ervan te overtuigen dat ze een goed produkt leveren: scholen zien vele opleidingsonderdelen, zoals theoretische componenten en het gestructureerd oefenen van vaardigheden als weinig zinvol. Wellicht moet deze ontwikkeling ook voor de Nederlandse opleiders een waarschuwing zijn. Het lijkt erop dat de politiek de neiging heeft inderdaad de opleidingsverantwoordelijkheid bij scholen te leggen, wanneer opleidingen onvoldoende tegemoet komen aan verlangens van de scholen of niet in staat zijn de onredelijkheid van die verlangens duidelijk te maken.

Onder de studies over opleidingsprogramma's viel ook dit jaar een presentatie van Klinzing en Floden op. Zij rapporteerden over een literatuurstudie naar de mogelijkheden om leraren te leren discussies te leiden. Daarbij kwamen ze tot een vorm van actie-onderzoek: de leraar experimenteert met zijn capaciteiten op dit gebied teneinde ze te verbeteren. Voor dat proces moeten leraren achtergrondkennis hebben, in staat zijn concepten te gebruiken om de analyse van hun gedrag te sturen en in staat zijn hypothesen te ontwikkelen. Ten slotte moeten ze de acties uit kunnen voeren om die hypothesen te toetsen en moeten ze van de resultaten van die toetsing kunnen leren. De auteurs gaven een overzicht van onderzoek naar de opleidingstechnieken op elk van die gebieden.

Minder technologische vormen van actie-onderzoek door leraren waren ook dit jaar nadrukkelijk aanwezig. Het ging in hoofdzaak om praktijkbeschrijvingen met weinig nieuwswaarde. De theoretische analyse van Noffke over de historie van actie-onderzoek vormde daarop een uitzondering.

Hoewel het aantal presentaties over reflectie wat afneemt is de belangstelling ervoor nog steeds groot. De conceptuele verwarring rond dit begrip duurt onverminderd voort. Betekenissen die men, veelal impliciet, aan reflecteren hecht zijn bijvoorbeeld nadenken, theoretiseren, het hanteren van metacognitieve strategieën, kritisch analyseren vanuit een historische context en leren als professional. In de bijdragen waarin men manieren presenteerde om studenten te leren reflecteren bleef het meestal bij een beschrijving van reflectie-activiteiten die door studenten in het opleidingsprogramma worden uitgevoerd. Het gaat dan bijvoorbeeld om logboeken bijhouden, actie-ondezoek doen, of om het uitvoeren van zelfstudieprojectjes. De laatstgenoemde activiteiten vertonen duidelijk verwantschap met de in Nederland door De Jong ontwikkelde en onderzochte explorations in de schoolstage. Aan deze studies lijkt de (aanvechtbare) aanname ten grondslag te liggen dat studenten vanzelf leren om zelfstandig te reflecteren wanneer ze maar veel moeten reflecteren. Veel van de verslagen over opleidingsprogramma's kwamen niet verder dan een globale beschrijving van een geheel programma, waarbij onduidelijk is wat nu precies de belangrijke elementen in het programma zijn. Korthagens presentatie van enkele concrete opleidingstechnieken stak hierbij gunstig af.

Handbook of Research on Teacher Education

Een hoogtepunt was ten slotte dat op de boekententoonstelling het *Handbook of Research on Teacher Education* (Ed.: W. R. Houston; uitgever: Macmillan) aanwezig was. Over dit standaardwerk en verschillende bijdragen daarin werd reeds geruime tijd door

ingewijden gesproken. De veelal goede bijdragen in dit boek (bijvoorbeeld van Floden & Buchmann en Carter) zijn nu voor iedereen beschikbaar.

Het onderzoek naar effectieve scholen (W. van de Grift, Rijksuniversiteit Utrecht)

Het belang dat aan het onderzoek naar effectieve scholen gehecht wordt komt mede tot uitdrukking in het aantal wetenschappelijke bijeenkomsten dat de Special Interest Group (SIG) 'School Effectiveness and School Improvement' kan organiseren op de AERA. Dit aantal bijeenkomsten, dat gebaseerd is op het ledental van een SIG, was beperkt tot drie, inclusief een ledenvergadering.

Rondom effectieve scholen zijn er nog steeds enkele hardnekkige misverstanden en misvattingen. Tijdens de lezingen op de AERA, maar vooral ook in de wandelgangen, is aan deze misvattingen aandacht besteed. Ik licht er een aantal punten uit.

Te vaak wordt nog de opvatting gehoord dat leerlingen van effectieve scholen hoge prestaties halen. Dat is natuurlijk niet zo maar waar. De aanhangers van deze opvatting zijn aanhangers van de koekenbakkertheorie (niet te verwarren met koekebakkertheorie, want dat betekent iets heel anders) over de effectiviteit van scholen. Een goede koekenbakker kan men kennen aan de hoge kwaliteit van zijn koeken. Het vakmanschap van een koekenbakker komt tot uitdrukking in het selecteren van de fijnste ingrediënten van de beste grondstof, en zijn vaardigheden in het bakproces. Het vakmanschap van een onderwijsgevende komt uitsluitend naar voren in het onderwijsproces. Een onderwijsgevende die de ingrediënten selecteert op milieu van herkomst en andere individuele kenmerken, zoals intelligentie en motivatie, discrimineert en kan voor de rechter gesleept worden. Juist de individuele kenmerken van leerlingen (zoals intelligentie, motivatie, milieu) verklaren echter zo'n slordige 80% van de variantie in leerlingprestaties. Dat geldt voor taalprestaties, dat geldt voor rekenprestaties en dat geldt voor prestaties in tal van andere vakken. Voor het beïnvloeden van de leerlingprestaties via het vakmanschap van onderwijsgevenden en hun schoolleiders is maar zo'n 20% van de variantie in leerlingprestaties beschikbaar. Een koekenbakker kan men echter gerust volledig op de kwaliteit van de koeken aanspreken. Een onderwijsgevende is pas echt goed als hij of zij naast de invloed die individuele leerlingkenmerken uitoefenen nog iets extra's heeft toe te voegen aan de leerlingprestaties.

Onderzoekers die scholen selecteren die hoog respectievelijk laag op een bepaalde test scoren en dan vervolgens gaan kijken naar de verschillen in schoolkenmerken, maken zich schuldig aan het aanhangen van de koekenbakkertheorie. Een voorbeeld daarvan is het onderzoek van Wellisch e.a. uit 1978. In 1978 kon men hiervoor nog een oogje dichtknijpen, omdat dit onderzoek een van de eerste meer empirische onderzoekingen was in reactie op de conclusies van met name Coleman, Jencks en Thorndike, dat scholen geen verschil maakten als het gaat om leerlingprestaties. Problematisch wordt het echter als mensen als Bamburg en Andrews dat nu in 1990 in Boston nog steeds doen. Juist zij waren bijvoorbeeld ook in Rotterdam met een presentatie, toen daar in 1989 het tweede congres over effectieve scholen gehouden werd. Kortom, misvatting nummer één, ook op de AERA, was nog steeds dat effectieve scholen scholen met hoge leerlingprestaties zijn. Effectieve scholen zijn scholen waarin leerlingen beter presteren dan op grond van hun milieu en intelligentie alléén verwacht mag worden.

Een tweede element dat een duidelijke ervaring in Boston te zien geeft, is het volgende. De periode van spectaculair onderzoek is voorbij. Het gaat niet langer om onderzoeksresultaten die het idee van een Kuhniaanse revolutie wekken. We zijn in een periode van 'normaal science' (Kuhn) terecht gekomen. Het is duidelijk geworden dat schoolfactoren maximaal zo'n 20% van de variantie in leerlingprestaties kunnen verklaren. De opdracht van de onderwijsonderzoekers is nu geworden om te gaan ontdekken welke van die factoren een plek hebben binnen die 20% te verklaren tussenschoolse variantie. Onder de Nederlandse onderzoekers op de AERA is het inmiddels gemeengoed dat dit ontdekkingsproces niet voltooid kan worden met eenvoudige middelen als eenvoudige correlationele berekeningen. Immers, als er maar ongeveer 20% variantie te verdelen valt, dan is de maximale multiple correlatie .45. Alle schoolfactoren (de attitude en het gedrag van de schoolleider, het didactisch handelen van de onderwijsgevenden, de onderwijsmethoden, de schoolorganisatie, enzovoort) tezamen kunnen dus

niet veel meer dan .45 met leerlingprestaties correleren. Dit heeft grote gevolgen voor het onderzoek naar geïsoleerde factoren, zoals bijvoorbeeld het gedrag van de schoolleider. De maximale correlaties die er voor afzonderlijke factoren, zoals de schoolleider, de onderwijsgevende, de schoolorganisatie, de onderwijsmethoden en dergelijke gevonden kunnen worden, zullen in de sfeer van de .10 en .20 liggen. De 'buit' moet immers met veel factoren gedeeld worden. Als men dan toch eenvoudig correlationeel onderzoek opzet naar slechts één of enkele factoren, dan is het noodzakelijk met repondentgroepen van minstens 400 scholen te werken. Als men dat niet doet, dan kan men bij de te verwachten correlaties in de sfeer van de .10 en de .20 nooit significante resultaten vinden. Dit zou wel eens de belangrijkste verklaring kunnen zijn voor het feit dat kenmerken van effectieve scholen die in het ene onderzoek gevonden worden, niet gereproduceerd kunnen worden in het volgende onderzoek.

Een derde element dat onder de onderzoekers die het thema effectieve scholen bezochten, een rol speelde, is het idee dat het niet verstandig is om slechts op één of enkele schoolkenmerken te jagen. Het gaat om het effect van een totale symfonie aan kenmerken. Vermoedens beginnen post te vatten dat bepaalde groepen van factoren tezamen bepaalde effecten oproepen, terwijl de afzonderlijke factoren weinig of geen effecten te zien geven. Het is in Boston meermalen betoogd: het moet gaan om het ontwikkelen van een theorie die ons iets leert over het specifieke samenspel van factoren dat effect sorteert. We hebben allemaal het gevoel: we kennen de afzonderlijke factoren, maar we weten niet in welke samenstelling ze effecten sorteren.

Beroepsonderwijs en bedrijfsopleidingen (W. J. Nijhof, Universiteit Twente)

Het is op de AERA niet gemakkelijk om de beide bovengenoemde trefwoorden binnen één sessie aan te treffen. Dat lukte ook dit keer niet. Bedrijfsopleidingen vormt evenals in Nederland een ander circuit dan de wereld van het beroepsonderwijs. De laatste – vocational education genoemd – heeft een ronduit slechte reputatie in het Amerikaanse onderwijssysteem. Toen in 1983 het beruchte rapport A Nation at Risk verscheen, samengesteld door een uiterst conservatieve commissie met hoofdzakelijk leden uit California en Texas, bleek met geen woord gerept te zijn over het beroepsonderwijs. Dat was aanleiding voor de American Vocational Association (AVA) onmiddellijk een eigen rapport uit te brengen in de kleur paars onder de titel The Unfinished Agenda. Research op het terrein van het beroepsonderwijs is dun gezaaid in de VS, alhoewel de AVA zelf een research divisie heeft. Er is veel meer ontwikkelingswerk.

Tijdens de AERA 1990 was opvallend dat er onderzoek werd gepresenteerd uit een vijftal instituten verspreid over de Verenigde Staten, die sinds 1988 gebundeld zijn tot het National Center for Research in Vocational Education, waarvan het hoofdkantoor zich in Berkeley, Stanford University bevindt. Deze situatie is daarom zo opmerkelijk, omdat zich voordien gedurende ruim 25 jaren een dergelijk Nationaal Centrum zich in Columbus, Ohio bevond, waar echter zelden of nooit onderzoek werd verricht. Door dat instituut is door de federale overheid een streep gezet.

Verwonderlijk aan het huidige onderzoek naar het beroepsonderwijs in de VS is de vraag naar de *integratie tussen beroepsonderwijs en algemeen vormend onderwijs*, terwijl vocational education al onderdeel is van de junior en senior high school.

Nog veel verwonderlijker is het dat Amerikanen van nul af beginnen, niet gehinderd door enige historische kennis van b.v. de Westduitse literatuur. Van Litt en Blankertz hebben zij nooit gehoord. De aanleiding tot die vraagstelling is uiteraard de veranderde arbeidsmarkt, de grote rol van technologie en de disfunctionaliteit en onderschatting van vocational education in de high school. De presentaties tijdens de AERA getuigden van descriptieve studies, beschrijvingen van experimentele modellen en voornemens om meer naar Europese en Japanse ontwikkelingen te kijken.

Evident is dat de vraag naar basisvaardigheden (transferable skills), die overigens in de jaren zeventig wel door het Centrum in Columbus, Ohio werd aangepakt, maar niet opgelost, nu weer als dominant vraagstuk op de agenda staat.

Het bovengenoemde consortium van onderzoeksinstellingen heeft een grootschalig 5-jarig onderzoeksprogramma opgesteld,

waarvan ook een longitudionale studie deel uitmaakt. Stern, en Stone en Hopkins resp. van de Universiteiten van California en Minnesota willen met name weten of er een relatie bestaat tussen werkervaringen en oriëntatie op werk. Veel Amerikaanse studenten hebben betaalde banen naast hun studie en dat levert problemen op voor de organisatie en planning van het onderwijs. Interessant is te weten of die werkervaringen ook positief doorwerken naar hun latere baan. De baseline gegevens dateren van 1988, de onderzoekers kunnen derhalve slechts voorlopige gegevens presenteren van 1 jaar. Uit hun eerste analyse blijkt dat er een sterke relatie bestaat met latere werkervaringen en succes op werk, los van elke vorm van schoolsupervisie. Als er banen zijn waarvoor de school supervisie geeft, vergelijkbaar met stages en ten dele met leerovereenkomsten in het leerlingstelsel, dan blijkt na controle voor demografische factoren en kenmerken van de baan, dat gesuperviseerde jobs significant lager scoren op motivatie en centrality (een maat voor het relatieve belang dat iemand toekent aan een baan). Betekenisvol is dat men tracht de baselinegegevens verder aan te vullen en indringende analyses te maken van het carrièreverloop.

Het is een vraagstelling die ook voor Nederland interessant kan worden indien vormen van open leren of community college concepten op grote schaal ingang zouden vinden.

Bedrijfsopleidingen

Evenals in Europa vormen bedrijfsopleidingen een belangrijke sleutel voor werknemers om een carrière te ontwikkelen. Het aanbod van papers en ronde-tafelsessies m.b.t. problemen op het gebied van bedrijfsopleidingen was echter uiterst sober: een sessie van de SIG Training in Business and Industry en twee ronde-tafelsessies. Twee werelden: AERA en bedrijfsopleidingen?

Tijdens de SIG werden in kwalitatief opzicht heterogene presentaties geboden, waarvan enkele betrekking hadden op een thematiek die ook in ons land een rol speelt op dit moment: effectiviteitsvraagstukken. Wolfe en Wolf, beide dames, waarvan de eerste hoofd opleidingen bij XEROX en de tweede hoogleraar aan de universiteit van South California, hebben een baseline ontwikkeld met data over elke door XEROX ontwikkelde en uitgevoerde cursus. Zij zijn nu in staat om via directe en secundaire analyses interessante vraagstellingen te beantwoorden. Het bleek evenwel niet eenvoudig een aantal van die vragen te presenteren. De voordracht beperkte zich tot een uitvoerige presentatie van het instrumentarium. Het lijkt de moeite waard om kennis te nemen van het instrument en van de wijze van gebruik daarvan, zoals terugkoppeling naar cursusbouwers en cursusmanagement.

Duffy (Xavier University of Louisiana) behandelde het bekende vraagstuk 'Wat is effectieve training?' Hij meent dat training slechts effectief is, wanneer de cultuur van een organisatie mensen motiveert tot het volgen van een training en het opdoen van kennis en vaardigheden. Daarnaast wijst hij uiteraard op specifieke trainingscondities die vaak van instructietechnische aard zijn. Hij deed een quasi-experimenteel onderzoekje naar twee instructievarianten, waarbij hij constateert dat er significante verschillen tussen beide treatments bestaan. De cursus (van 12 uren) ging over een introductie op het gebied van microcomputers. Duffy schrijft de verschillen, gemeten via een Likerttype satisfactie-schaal, (8 items), toe aan de grotere mate van bereidheid van organisatie B om mensen aan te zetten tot training. Dit onderzoek bewijst dat er veel te verbeteren valt op het gebied van het meten van effecten van bedrijfsopleidingen. Niet alles wat de AERA te bieden heeft kan de toets der kritiek doorstaan.

Gustafson e.a. deden een onderzoek naar de effecten van een oriëntatie- en trainingsprogramma m.b.v. interactieve videodisc bij de firma Apple. In dit onderzoek is getracht korte en lange termijn effecten te meten en is ook nagegaan hoe de cursusopzet verbeterd zou kunnen worden. Gebruik is gemaakt van een posttest, een vragenlijst en een gedragstest, waarin de nieuwe vaardigheden moesten worden getoond, om de korte termijn effecten te meten. De resultaten daarvan waren positief. De lange termijn effecten zijn telefonisch bij betrokkenen verkregen na 6 tot 8 weken. Veel van de effecten zijn echter vermoedelijk toe te schrijven aan de training tijdens het werk na de cursus. Echte controle daarop is niet uitgevoerd.

Ook hier een voorbeeld van effectiviteitsmetingen en cursusverbetering, waarbij de controle op design en meetinstrumentarium beter zou kunnen.

Interessant was het daarom zeker om bij een roundtable sessie te beluisteren wat William Ives te brengen had op het gebied van de methodologie van kwaliteits- of deskundigheidsbevordering. Volgens Ives is excellent gedrag te definiëren in kwantitatieve en kwalitatieve termen en is het gebonden aan de missie van een onderneming, aan zijn strategie en aan zijn richting. Hij ontwikkelde daartoe een 'performance system development model' dat qua gedachtengang sterk gebaseerd is op het werk van Deming dat uitgaat van een analyse van excellent gedrag van de werknemer. Ives' 6-fasenmodel is een waardevolle bijdrage om het werk van en in een bedrijfsopleiding in kaart te brengen en bevat instrumenten om het qua architectuur nader te ontwikkelen

Hopelijk neemt niet alleen het aantal bijdragen op dit domein toe, maar ook de kwaliteit ervan. Om dat te bevorderen is een researchagenda nodig. Mulder (Universiteit Twente) presenteerde uit het onderzoek Bedrijfsopleidingen in Nederland de resultaten van een panelstudie naar een dergelijke researchagenda. De belangstelling voor de methodiek van de panelstudie, zoals die door de vakgroep Curriculumtechnologie van de Universiteit Twente was gevolgd was groot. De vraag blijft wel hoe bepaalde typen vragen zich verhouden gegeven de verschillen in cultuur per land en per bedrijf.

Postsecundair onderwijs (J. F. M. J. van Hout, Katholieke Universiteit Nijmegen)

Divisie J (Postsecondary education) organiseerde in Boston 45 eigen sessies, waarvan 15 symposia en 25 papersessies (ca 125 papers). Deze divisie verzorgde daarmee ca 10% van de conferentie en is naar dit criterium gemeten een middelgrote divisie.

Researchoriëntatie

De meeste papers hebben betrekking op 'students, faculty/administrators, teaching and learning'. In het verslag van het AERA-congres van 1989 heb ik vooral aandacht besteed aan het onderzoek naar *faculty and staff*, omdat omvangrijke landelijke survey-onderzoeken van NCRIPTAL (National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan) en van OERI (Office of Educational Research and Improvement, U.S. Department of Education) werden aangekondigd. Dit jaar werden hiervan enkele voorlopige resultaten gepresenteerd. Verwacht mag worden, dat binnenkort de eerste artikelen hierover in de Amerikaanse tijdschriften over hoger onderwijs zullen verschijnen (bijv. "Research in Higher Education", "The Journal of Higher Education" en "The Review of Higher Education").

Uit het onderzoek van Blackburn (NCRIPTAL) dat in dit kader werd verricht, blijkt het reeds eerder gevonden beeld (o.a. Finkelstein, 1984) bevestigd te worden: stafleden van de top-universiteiten zijn veel sterker researchgeoriënteerd en minder onderwijsgeoriënteerd dan die van de andere universiteiten. De verschillen in aantallen publikaties tussen stafleden van universiteiten worden, althans volgens de voorlopige resultaten van dit grote survey, beter voorspeld door attitude-aspecten en de beschikbaarheid van 'grants', dan door achtergrondkenmerken (zoals leeftijd en geslacht). Ashburn en Peng presenteerden eerste gegevens van een nationaal survey (OERI), op basis waarvan zij concluderen, dat het zwakke research-imago van de 'Education Faculty' niet zozeer een gevolg is van significant lagere produktiviteit (publikaties), maar vooral van een geringere tijdsbesteding aan research en dus wellicht een andere attitude ten aanzien van research.

Cursusplanning

Bijzondere aandacht verdient het onderzoek van Stark en Lowther (NCRIPTAL) over *cursusplanning* door wetenschappelijke staf, verbonden aan Colleges. Naar mijn weten is dit het eerste grote onderzoek naar opvattingen en gedrag van stafleden op het gebied van cursus- en curriculumplanning op postsecundair niveau. Uit de eerste resultaten blijkt, dat de wetenschappelijke staf nauwelijks over theoretische kennis beschikt met betrekking tot curriculumconstructie en -technologie. Zij laten zich ook nauwelijks door experts op dat gebied adviseren. De meeste invloed op hun denken hebben de eigen praktische ervaringen, de kenmerken van de vakinhoud en de collega's c.q. de traditie. Er blijken duidelijke verschillen tussen de disciplines te bestaan. Dit geldt niet alleen ten aanzien van de opvattingen, maar ook ten aanzien van het gedrag, zoals de frequentie waarmee cursussen worden gereviseerd, dan wel opnieuw opgezet. Hoewel degenen die zich professioneel met

begeleiding van docenten in het hoger onderwijs bezighouden door deze gegevens niet verrast zullen zijn, ben ik van mening dat dit onderzoek uitermate relevant is. Kwaliteit van onderwijs is sterk afhankelijk van de wijze waarop het onderwijs is opgezet. Dit is vooral zo, indien gedacht wordt aan herstructurering van onderwijs volgens principes van docent onafhankelijk onderwijs, modularisering e.d. De gemiddelde docent in het hoger onderwijs heeft niet geleerd vanuit een onderwijskundig referentiekader naar het eigen onderwijs en dat van andere docenten te kijken, en op grond van die analyse aan de systematische verbetering ervan te werken. In het postsecundair onderwijs wordt cursusoptimalisering vooral vanuit aanpassing van de inhoud, de leerstof, aangepakt, zoals ook uit het onderzoek van Stark en Lowther blijkt. Door het werk van deze onderzoekers wordt een empirische basis gegeven aan zaken, die nu nog te gemakkelijk afgedaan (kunnen) worden als vooroordelen van belanghebbende professionals. Een dergelijk onderzoek zou daarom ook voor Nederland verfrissend kunnen werken.

Dropout

Ieder jaar worden ook enkele papers gepresenteerd over de problematiek van *dropout* ofwel *student attrition*. Door Tinto is aan dit thema eerder een theoretische basis gegeven, die tot cumulatie van kennis en een aparte stroom van onderzoek heeft geleid. Desalniettemin is de 'persistence rate' in de U.S. voor colleges en universiteiten volgens Tinto gedurende de periode 1880-1980 onveranderd gebleven op ca 52%. Van de 2.65 miljoen studenten die zich in 1989/90 voor het eerst inschreven, zal ca 2.58 miljoen niet afstuderen bij de instelling van eerste inschrijving en ca 1.15 miljoen studenten geen enkel diploma behalen, aldus Melvin Webb II van het ACTP (American College Testing Program). Aan de maatschappelijke relevantie van dit onderzoek hoeft dus niet getwijfeld te worden, aan het effect op de dropout tot nu toe wellicht des te meer. De discussie richt zich daarom op de bruikbaarheid van Tinto's model en het daarop geënte onderzoek voor het hele (zeer diverse) Amerikaanse hoger onderwijs. Veel onderzoek beperkt zich feitelijk tot een bepaald type (de 4-jarige 'residential' opleiding), terwijl het 'commuter college' steeds belangrijker wordt vanwege deelname van zgn. niet-traditionele studenten-groepen. Naast het Tinto-model zijn daarom varianten ontwikkeld door Pascarella en door Bean, waarin de sociale integratie (typerend voor de campus-universiteit) een minder belangrijke rol speelt.

Melvin W. Webb II ontwikkelde een aangepast causaal model voor verklaring van uitval bij commuter instellingen en toetste dit model op een aantal commuter en residential instellingen: bij de commuter instellingen bleek de sociale integratie inderdaad geen effect te hebben, terwijl deze factor wel een sterk direct effect op uitval heeft bij residential colleges. Dit resultaat is ook van belang voor de Nederlandse onderzoekers, omdat het Nederlandse hoger onderwijs evenmin gekenmerkt wordt door een sterke sociale integratie van studenten in de onderwijsorganisatie en het studentenleven: het door Webb ontwikkelde model zou daarom wellicht beter aansluiten bij de kenmerken van ons hoger onderwijssysteem dan het Tinto-model.

Bean en Vesper presenteerden een onderzoek waarin een model werd ontwikkeld voor een aparte instelling. Zij wilden hiermee aantonen, dat met behulp van een Lisrel-benadering lokale theorieën ontwikkeld kunnen worden die hoog scoren op beleidsrelevantie en bovendien methodologisch veel meer waarde hebben dan kwalitatief onderzoek. Een heel opmerkelijk gegeven is, dat de 'intention to leave' die in eerdere correlatie-studies een belangrijke rol speelt, als causale factor volledig blijkt te verdwijnen. De invloed van ouders (instemming met/afkeuring van vertrek) blijkt een zeer belangrijke variabele te zijn. Deze aanpak van Bean en Vesper lijkt mij zeer waardevol, omdat ik denk dat het rendementsprobleem moeilijk hanteerbaar is op landelijk niveau, afgezien van structurele maatregelen betreffende de cursusduur en de inschrijvingsduur. Er zijn in het bestaande systeem verschillen in numeriek rendement tussen disciplines en tussen instellingen, maar dat betekent nog niet dat – bestuurlijk gezien – op dat niveau het meest effectief maatregelen ter verbetering van het rendement genomen kunnen worden. Naar mijn mening zal vooral het niveau van de studierichting en de faculteit effectieve maatregelen kunnen nemen op het vlak van selectie, opzet van het onderwijs, kwaliteit van toetsing etc. ter verbetering van het rendement. De aanpak van Bean en Vesper kan leiden tot studierichting-

specifieke modellen ter verklaring van uitval en vertraging, die zowel recht doen aan de complexiteit van de problematiek als voldoende handvatten aanreiken voor verantwoordelijke bestuurders.

Toetsen en meten (K. Bügel, H. Robben, D. Thio en P. Sanders, CITO, Arnhem)

Presidential Address NCME (National Council on Measurement in Education)
Dit jaar sprak Ronald Hambleton de jaarlijkse voorzittersrede uit. Hij merkte op dat het voor 'educational measurement' goede en kwade tijden zijn. Er is de afgelopen 25 jaar heel veel technisch-wetenschappelijke vooruitgang geboekt; anderzijds staan testmensen en de professie als geheel heden ten dage aan veel kritiek bloot. In de nabije toekomst ziet Hambleton criterion referenced testing, higher order thinking skills en IRT-modellen en adaptief toetsen belangrijker worden. De invloed van IRT op ook praktische testprocedures is langzamerhand op een breed front merkbaar. Computers en nieuwe media zullen adaptief toetsen steeds meer realiseerbaar maken. Het aantal banen voor testdeskundigen is dan ook groot, het (studenten)aanbod echter gering. De scholing van docenten met betrekking tot toetsen en evalueren schiet nog altijd erg tekort. Recent hebben de NCME en twee lerarenorganisaties besloten samen te werken bij de opstelling van een programma van eisen: 'wat moet een docent van educational measurement weten?'

Multilevel software
Het feit dat leer- en onderwijsprocessen plaatsvinden binnen klassen en scholen, is aanleiding geweest tot de introductie en ontwikkeling van multilevel technieken in de onderwijsresearch. Namens een onderzoeksgroep van UCLA deed Kreft verslag van een onderzoek waarin multilevel software, met name de programma's GENMOD, HLM, ML2 en VARCL, werd vergeleken. De algemene conclusie was dat hoewel alle programma's goed te gebruiken zijn ze soms nogal gecompliceerd kunnen zijn. De gebruiker moet vooral uitkijken bij kleine datasets en wanneer er veel parameters geschat moeten worden. In zijn reactie besprak Prosser van het Institute of Education uit Londen de sterke en zwakke punten van ML3. Hij wees ook op de mogelijkheid tot het volgen van cursussen die regelmatig in Londen gegeven worden. Nieuwe ontwikkelingen betreffen o.a. de ontwikkeling van een computerprogramma, genaamd MULTIPATH door De Leeuw, Kreft, en Kim, en een begin volgend jaar te verschijnen boek over hiërarchische lineaire modellen door Bryk en Raudenbush.

Testtheorie
Veel lezingen hadden betrekking op ontwikkelingen maar vooral toepassingen van itemresponstheorie. Met name de SIG Rasch Measurement is elk jaar duidelijk aanwezig op de AERA. Wat betreft itemresponstheorie verkeren Nederlanders echter in de gelukkige omstandigheid dat nieuwe en interessante ontwikkelingen op dit gebied eerder in Nederland dan in de V.S. plaatsvinden. Zo hebben psychometrici van het Cito de afgelopen jaren een generalisatie (o.a. polytome data en geïmputeerde discriminatie parameters) van het Rasch model ontwikkeld. De software voor dit model is binnenkort verkrijgbaar.

Verrassend was het aantal lezingen (12) dat de generaliseerbaarheidstheorie tot onderwerp had. Een sessie met vier lezingen betrof het gebruik van de generaliseerbaarheidstheorie bij de analyse van mondelinge examens. Naast een sessie met zowel univariate als multivariate toepassingen, was er ook een theoretische sessie. Luecht en Smith bespraken het effect van 'unbalanced data' op de schatting van variantiecomponenten. Op basis van een vergelijking van negen methoden, o.a. ANOVA en (restricted) maximum likelihood, was hun conclusie dat voor de door hen onderzochte eenvoudige designs de verschillen tussen de methoden gering waren en dat niet één methode als de beste kon worden aangemerkt. Het paper van Smith had betrekking op de invloed van gecorreleerde effecten op de schattingen van de variantiecomponenten. Zijn voorlopige conclusie was dat men zich het beste tegen het optreden van bias kon beschermen door bij een facet meer dan tien condities te gebruiken.

Computer Adaptief Toetsen (CAT)
CAT is vooral mogelijk geworden door toepassing van itemresponstheorie. CAT is een aantrekkelijk alternatief voor traditionele paper-and-pencil toetsen omdat het met CAT mogelijk is leerlingen met minder items dan op de traditionele manier even betrouwbaar te

meten. Koch en Dodd bespraken de effecten van de informatiefunctie van de itemverzameling op CAT wanneer het partial credit model gebruikt wordt. In eerder onderzoek hadden zij geconstateerd dat itemverzamelingen van slechts 30 items voldoende waren om vaardigheidsschattingen met een gewenste nauwkeurigheid (stopregel) te verkrijgen. Deze itemverzamelingen hadden echter symmetrische, gepiekte informatiefuncties. De resultaten van hun onderhavige onderzoek wezen uit dat itemverzamelingen met scheve informatiefuncties tot problemen aanleiding geven en dat items aan de itemverzameling toegevoegd moeten worden teneinde de scheefheid te reduceren. In het kader van CAT vergeleken Dodd en Koch twee stopregels, nl. de minimum standaardfout stopregel en de minimum informatie stopregel. Zij concluderen dat voor itemverzamelingen bestaande uit 30 partial credit items de twee stopregels vergelijkbaar zijn, maar dat bij itemverzamelingen bestaande uit meer dan 60 partial credit items de minimum standaardfout stopregel de voorkeur verdient, d.w.z. leidt tot dezelfde meetnauwkeurigheid met minder items.

Meten van veranderingen

Muthén liet zien hoe veranderingen van responses op items als gevolg van instructie (Opportunity To Learn) gemodelleerd kunnen worden met het LISCOMP model. Embretson stelde een multidimensioneel Rasch model voor om veranderingen van leerlingen te kunnen kwantificeren. Wilson presenteerde uitbreidingen van het partial credit model die het mogelijk maken informatie over diverse oplossingsstrategieën te verkrijgen.

Itembias en sekseverschillen

Aan de detectie van itembias en sekseverschillen in toetsprestaties waren ongeveer 30 lezingen gewijd. De lezingen konden onderscheiden worden in:

1. Lezingen die de Mantel-Haenszel itembias procedure tot onderwerp hadden;
2. Lezingen die betrekking hadden op andere itembias procedures;
3. Lezingen die inhoudelijke verklaringen van itembias betroffen;
4. Lezingen die betrekking hadden op sekseverschillen in toetsprestaties.

Ad 1

Clauser wees op het verschijnsel dat wanneer een toets met heterogene items wordt opgedeeld in homogene subtoetsen er niet alleen meer maar ook gedeeltelijk andere items met itembias gedetecteerd worden. DeMauro liet zien dat de MH-procedure gevoelig is voor de verhouding van de kandidaten in de te onderzoeken groep, bijv. zwarten, en de referentie groep, bijv. blanken. Hij benadrukte het belang van representatieve steekproeven.

Ad 2

Rogers vergeleek een logistische regressie methode voor de detectie van itembias met de MH-procedure en concludeerde dat haar methode ongeveer dezelfde resultaten opleverde maar dat de MH-methode goedkoper was. Seong verdedigde de opvatting dat variantie analyse een geschikte techniek is voor het opsporen van itembias bij kleine groepen kandidaten. Andere itembias procedures waren gebaseerd op itemresponstheorie. Ferrarra liet zien hoe bias bij toetsen met open vragen ('partial credit model') opgespoord kan worden. Camilli was van mening dat itembias het beste gezien kan worden in het licht van een multidimensioneel itemrespons model. Itembias wordt zijns inziens veroorzaakt door verschillen in vaardigheid op een secundaire latente trek, die onbedoeld ook gemeten wordt.

Ad 3

Verschillen in voorkennis zouden volgens Scheuneman verklaren waarom vrouwen op sommige tekstbegriptoetsen gemiddeld lagere prestaties behalen dan mannen. (Een opvatting die aansluit bij de bevindingen van onderzoek dat op dit moment op het Cito uitgevoerd wordt.) Verschillen worden vooral veroorzaakt door tekstbegriptoetsen met een natuurwetenschappelijke inhoud. Vrouwelijke biologiestudenten met meer voorkennis behaalden namelijk dezelfde resultaten als mannelijke biologiestudenten. Lunz e.a. vonden dat er grote verschillen tussen beoordelaars bestaan wanneer zij verzocht worden itembias te verklaren.

Ad 4

Veel van de lezingen over sekseverschillen bij toets- en schoolprestaties hadden betrekking op het vak wiskunde. Op basis van een meta-analyse op 100 studies concludeerde Hyde dat in de V.S. de prestaties tussen mannen en vrouwen verschillen. Smith liet zien dat de

prestaties van meisjes op de Scholastic Aptitude Test Mathematics (SAT-M) achter blijven bij die van jongens. Verklaringen hiervoor werden in andere lezingen gesuggereerd. Fry meende dat de verschillen veroorzaakt worden door andere oplosstrategieën van meisjes. Meisjes passen vaker op school geleerde regels en strategieën toe, terwijl jongens meer gokken en meer globaal te werk gaan, wat bij de SAT betere resultaten oplevert. Uit ander onderzoek bleek dat meisjes betere resultaten behaalden dan jongens op enkele opgaven uit de SAT wanneer de tijdsdruk verdween. Uit een protocolanalyse van Levine bleek dat meisjes banger waren om fouten te maken en minder bereid waren problemen op een onconventionele wijze op te lossen. Uit verschillende onderzoeken bleek dat ook het type toets invloed heeft. Vrouwen raden minder dan mannen bij meerkeuze items en ze maken vaker gebruik maken van het antwoord 'Ik weet het niet'.

Aan deze Kroniek werkten mee: *E. van den Berg, G. J. J. Biesta, K. Bügel, W. van de Grift, J. F. M. J. van Hout, L. F. W. de Klerk, S. Miedema, M. Mulder, W. J. Nijhof, H. Robben, P. Sanders, D. Thio, Th. Wubbels*

De eindredactie werd verzorgd door
W. J. Nijhof

Boekbesprekingen

Slavenburg, J.H. & T.A. Peters (red.), *Het project Onderwijs en Sociaal Milieu: een eindbalans*. Rotterdamse School Advies Dienst, Rotterdam, 1989, 380 pag., *f* 35,–.

Gedurende 20 jaar heeft zich in Rotterdam een onderzoeksproject afgespeeld dat waarschijnlijk het omvangrijkste en het duurste (bijna 50 miljoen!) was uit de geschiedenis van het Nederlandse sociaalwetenschappelijk onderzoek. Einde 1989 is het afgesloten, ter gelegenheid waarvan de onderzoekers de balans hebben opgemaakt.

Een van de merkwaardigheden van dit zo praktijkgerichte onderzoeksproject is altijd geweest de ruimte die er bestond voor instrumenteel en anderszins fundamenteel onderzoek. Talloos waren de papers over methoden, technieken, operationalisaties e.d. waarvan het belang uitging boven de directe locale bruikbaarheid. Lof verdient de projectleiding, die, dat blijkt ook weer uit deze eindbalans, tijd en ruimte heeft weten te vinden om de internationale literatuur kritisch te volgen.

Tijdens de afsluitingsdag kwam nog eens helder aan het licht dat de onderzoekers hebben geprobeerd vanuit een methodologisch zuiver standpunt hun werk te doen. Slavenburg, opvolger van initiator Grandia en belangrijkste projectleider, plaatste de vooraf geformuleerde doeleinden van het project nog eens helder voor het voetlicht, en constateerde op grond van de onderzoeksresultaten dat de meeste doeleinden niet gehaald zijn. In die zin moet het project dan ook als mislukt worden aangemerkt. Op hem volgende sprekers, waaronder de huidige bewindslieden Simons en Wallage, deden vervolgens hun best om duidelijk te maken dat er toch veel bereikt is, dat 'velen er toch veel van geleerd hebben', enz. Niet alleen deze toehoorder voelde zich wat ongemakkelijk bij het merkwaardige dilemma van de 'objectieve mislukking' en de kennelijke behoefte van beleidszijde om het project positief te waarderen.

In de eindbundel is in het kort de geschiedenis van het O.S.M. project uiteengezet, waarbij de veelzijdigheid – het werken met leerlingen, leerkrachten, materiaal, ouders, aanvankelijk ook buurten – opvalt. In verschillende hoofdstukken komen de ervaringen met curriculumontwikkeling, activerings- en onderwijsbegeleidingsprogramma's en met de O.S.M. verspreidingsstrategie aan de orde. Ook is een hoofdstuk gewijd aan buitenlandse en andere Nederlandse stimuleringsprogramma's. Aparte beschouwingen krijgen de relaties tussen leerlingkenmerken, gezinskenmerken en schoolprestaties. Het hoofdstuk over leerlingkenmerken en schoolprestaties laat helaas veel vragen onbeantwoord betreffende de rol van de intelligentie ten opzichte van schoolvorderingentoetsen. De auteurs verdedigen de gekozen en veelal niet bereikte 'transfercriteria', waarmee bedoeld wordt dat de eisen gesteld werden in relatie tot landelijke gemiddelden (p. 106 e.v.). Deze criteria hebben betrekking op lezen, rekenen, zittenblijven, de uitstroom naar het speciaal onderwijs enerzijds en naar het voortgezet onderwijs anderzijds. In de slotbeschouwing wordt overigens erkend dat het beter ware geweest om 'relatieve verbetering' als einddoel te stellen in plaats van het te ambitieus gebleken 'gemiddelde Nederlandse niveau'.

Hoewel deze bundel hier en daar moeilijk toegankelijk is – in de loop van 20 jaar heeft zich uiteraard een eigen denkwijze en jargon ontwikkeld – kan men er veel uit leren voor wat betreft het plannen en uitvoeren van een evaluatieonderzoek. De onderzoekers verdienen grote lof voor de zorgvuldige explicitering van hun eigen werk en voor hun onbevangenheid terzake van het niet bereiken van de gestelde doeleinden; voor de vanzelfsprekendheid ook waarmee zij er van uitgaan dat kostbare begeleidingsprojecten aanwijsbare resultaten *moeten* hebben. Voor wat dit laatste punt betreft is er geen groter contrast denkbaar dan met de recente dissertatie van Donker (1990) over evaluatie van projecten in de geestelijke gezondheidszorg; na het niet kunnen vaststellen van effecten constateert de auteur dat de meeste projecten – en het bijbehorende onderzoek – toch zin hebben, al was het alleen maar omdat de voortdurende onderzoeksbegeleiding de projectmedewerkers dwingt tot een permanente reflectie.

Voor een evaluatie van specifieke projectdoelen en -aspecten die hout snijdt heeft de lezer echter méér nodig dan alleen deze publikatie. Verwacht – en gehoopt – mag worden dat het laatste woord over de wijze waarop men in Rotterdam het probleem van de ongelijke onderwijskansen heeft aangepakt en geëvalueerd nog niet gezegd is.

P. G. Swanborn

Literatuur

Donker, M., *Principes en practijk van programma-evaluaties*. Utrecht: Nederlands Centrum Geestelijke Volksgezondheid, 1990.

R. A. M. van Brunschot *Begeleiding van beginnende docenten in het technisch onderwijs*. Proefschrift Eindhoven, 1990, 284 pag. + 12 pag. bijlagen, *f* 46,–, ISBN 90 9003256 8.

Het boek doet verslag van onderzoek naar problemen van beginnende docenten in het technisch onderwijs en de effecten van begeleiding van beginnende docenten in het middelbaar technisch onderwijs. Van Brunschot spitst dit toe tot drie onderzoeksvragen die hij aldus verwoordt (p. 16):

'1. Bestaan er verschillen tussen de probleemervaring van beginnende docenten die werkzaam zijn in het Basisonderwijs en Voortgezet Onderwijs enerzijds en in het Hoger Beroeps Onderwijs anderzijds?
2. Bevat de literatuur op het gebied van het opleiden van docenten aanwijzingen dat een begeleiding, gericht op het leren oplossen van probleemsituaties, een zinvolle manier is om beginnende docenten te helpen?
3. Heeft een begeleiding gericht op het leren oplossen van probleemsituaties het beoogde effect: zijn beginnende docenten als gevolg van die begeleiding in staat problemen die zij als docent ervaren, op een meer systematische manier aan te pakken en verbetert daardoor hun functioneren?'

Het antwoord op de eerste vraag wordt gezocht in een inventarisatie van probleemervaringen van beginnende docenten, dat wil zeggen docenten met minder dan twee jaar leservaring (p. 38). Daartoe worden 58 beginnende HBO-docenten en 26 beginnende MTO-docenten bevraagd naar hun ervaringen, waarna de resultaten worden vergeleken met die van andere studies onder beginnende docenten in basisonderwijs en voortgezet onderwijs. Geconcludeerd wordt dat beginnende HBO-docenten problemen ervaren op didactische aspecten, terwijl BO- en VO-docenten vooral op relationele aspecten van docentgedrag problemen ervaren. MTO-docenten tenderen een middenpositie in te nemen (p. 54).

Als antwoord op de tweede vraag wordt een overzicht van onderwijsvisies gegeven waarna een begeleidingsmodel wordt gekozen, namelijk het probleemaanpak-model van Egan. De auteur geeft vervolgens een uitvoerige beschrijving van dit model.

De beantwoording van de derde vraag, de hoofdmoot, wordt gezocht in een onderzoek naar effecten van het modelgebruik. Negen ervaren MTO-docenten worden geoefend in het gebruik van Egans model, waarna zeven van hen in totaal elf beginnende docenten volgens dat model begeleiden. Met behulp van een vragenlijst wordt hun mening omtrent het gebruik van het model tijdens de begeleiding gepeild. Bovendien wordt getracht het effect van de begeleiding op beginnende docenten te bepalen door hun een video-opname van één klassesituatie te vertonen met als opdracht hardop denkend aan te geven hoe men in die probleemsituatie zou reageren. Naast de begeleide beginners die de experimentele groep vormen, zijn er onbegeleide beginners die de controle groep vormen. Tot slot wordt met behulp van een vragenlijst onderzocht of er naar de mening van betrokkenen ook sprake is van een verbetering in hun functioneren. Geconcludeerd wordt dat de begeleiding heeft geleid tot meer nadenken over de probleemsituatie (p. 216) en dat het al of niet begeleiden geen significant verschil oplevert in de beoordeling van het eigen functioneren (p. 230).

Over het algemeen is de tekst helder geschreven en goed gestructureerd. Het inleidende hoofdstuk geeft een duidelijk overzicht van het onderzoek, terwijl in het slothoofdstuk de bevindingen in samenhang op een rijtje worden gezet. Het streven om hoofdstukken zo te schrijven dat ze los van elkaar te lezen zijn (p. 7), heeft tot veel herhalingen geleid.

Ten aanzien van de keuze van de probleemoplossingsbenadering en in het bijzonder de keuze van het model van Egan valt op dat de gegeven argumentatie wel erg summier is. Het belang van het model in het onderzoek rechtvaardigt ongetwijfeld de uitvoerige beschrijving die Van Brunschot ervan geeft (circa 25 pagina's). Maar om dezelfde reden zou een gedegener verantwoording van de gemaakte keuze mogen worden verwacht, zeker gelet op de diversiteit aan beschikbare modellen zoals ook Van Brunschot signaleert (pp. 111-112).

Vooral in het hoofdonderzoek is het empirisch materiaal dat wordt gepresenteerd, in meerdere opzichten schraal. De begeleiding wordt beproefd door zeven begeleiders bij in totaal 11 beginners. Met behulp van vragenlijsten worden vervolgens hun meningen gepeild. Er wordt dus niet onderzocht of beginners door begeleiding beter zijn gaan functioneren, maar of zij zelf denken beter te zijn gaan functioneren, waarbij voorts onduidelijk blijft welke normen individuele docenten hanteren. Dit is ook het geval bij de video-opname. Deze en andere bezwaren van methodologische aard worden overigens wel door de auteur onderkend (zie bijvoorbeeld pp. 139-144, p. 167 en p. 224).

De vraag of begeleiding van beginnende docenten tot een daadwerkelijke verbetering in hun functioneren leidt, wordt niet beantwoord en kan ook niet beantwoord worden met de door Van Brunschot gevolgde meetmethode van de 'zelfobservatie' (p. 141). De bijdrage die dit boek levert aan de discussie over effecten van begeleiding, moet dan ook beperkt worden geacht.

W. Jochems

Tomic, W. & P. C. van der Sijde, *Changing teaching for better learning*, Swets & Zeitlinger, Amsterdam, 1989, 103 pag., *f* 29,95, ISBN 90 265 1019 5.

Welk leerkrachtgedrag leidt eigenlijk tot leerresultaten? en: Hoe kunnen leerlingprestaties en attitudes verbeterd worden? Deze vragen gelden als 'leitmotiv' van een Engelstalig boek, waarin twee Nederlandse onderzoeken bijeen zijn gebracht. Het gaat om onderzoek dat verricht is in het kader van 'The Classroom Environment Study': een project onder auspiciën van de International Associaton for the Evaluation of Educational Achievement (IEA), uitgevoerd aan de Universiteit Twente.

De vragen verraden al direct dat we te maken hebben met de (rijke) traditie van leerkrachteffectiviteitsonderzoek. Hoewel de auteurs niet direct met zoveel woorden toegeven tot deze traditie te behoren, blijkt uit de opzet van hun onderzoek dat zij behoren tot de 'hard-boiled' variant binnen de traditie, gekenmerkt door een sterk empirisch (in plaats van theoretisch) georiënteerde aanpak. Het traject dat dergelijk onderzoek volgt, mag bekend verondersteld worden: men zoekt leerkrachtgedragingen (in een bepaald domein) die manipuleerbaar zijn en observeerbaar gemaakt kunnen worden om deze op samenhang met leerprestaties te onderzoeken.

De beschrijving van de correlationale studie is een heldere, compacte en sterk op de onderzoeksuitvoering georiënteerde uiteenzetting. De lezer wordt meegenomen langs een aantal in het onderzoek genomen beslissingen en de verantwoording ervan. Het draaiboek van een proces-produktstudie wordt correct gevolgd, wat ondermeer inhoudt dat uitvoerig aandacht besteed wordt aan het voorkomen van direct geobserveerde gedragingen en composieten daarvan. De resultaten wijken niet af van wat in de Amerikaanse proces-produktstudies is gevonden: Veel frontale instructie door de leerkracht, gericht op groepen leerlingen. Dit frontale onderwijs, zo concluderen de auteurs, is dominant omdat het het beste compromis is voor de dilemma's waarmee leraren geconfronteerd worden. In tegenstelling tot Amerikaanse studies zijn geen noemenswaardige effecten op leerprestaties gevonden (enkele kleine effecten van infrequente gedragingen daargelaten). De uitkomsten zijn verkregen middels talrijke afzonderlijke regressie-analyses met een steeds wisselende volgorde van predictorvariabelen.

Na de correlationale fase komt een belangrijk omslagpunt in proces-produktonderzoek (Fenstermacher, 1981). Effectieve gedragingen dienen geselecteerd te worden om te zien of via training deze ook verbetering van leeruitkomsten veroorzaken. De studie van Tomic en Van der Sijde gaat hier weinig klassiek te werk. Eigenlijk (p. 29) ontkennen zij de centrale 'improvement' hypothese in proces-produkt-onderzoek en stellen dat uit een samenhang tussen leerkrachtgedrag en leerresultaten niet geconcludeerd hoeft te worden dat dergelijk leerkrachtgedrag voor training in aanmerking komt. Hoe verdedigbaar de stelling ook mag zijn, het haalt de basis weg van de door de auteurs zo centraal gestelde: observational-correlational-experimental loop. Gevolg is dat de auteurs met lege handen staan voor wat betreft de invulling van de trainingsfase. Uit nood vallen ze terug op de observatie dat nogal (?) veel tijd (10%) is besteed aan niet-leerstofgerichte activiteiten wat mogelijkerwijs zou kunnen duiden op gebrek aan managementvaardigheid bij de leerkracht. Al met al een weinig onderbouwde, maar wel cruciale 'switch' in het onderzoeksdesign.

De trainingsstudie is qua theoretisch gedachtengoed met wat meer omhaal van woorden beschreven dan de correlationele studie. De legitimering van het cognitieve model van klasse-instructie en management is echter niet het sterkste deel uit het boek. Een instructiemodel wordt aanbevolen met daarin een aantal vaste stappen waarlangs de lesuitvoering dient te verlopen: controle van huiswerkpresentatie van lesstof, gecontroleerde oefening, begeleide oefening en huiswerk. Terecht stellen de auteurs, verwijzend naar Amerikaans onderzoek, dat in het algemeen te weinig aandacht is besteed aan de implementatie van trainingen. De auteurs specificeren in de try-out welke tijdsindicatie aan de lesstappen moet worden besteed (als normering van gedragsveranderingen in de training).

De implementatiestudie wijst uit dat de training inderdaad de tijdsverdeling van leraren meer in de gewenste richting doet veranderen.

Rest de laatste stap: leidt getraind gedrag tot betere leerresultaten? De auteurs volgen hier niet het design uit proces-produktstudies, maar hanteren een eigen uitbreiding ervan: vergelijking tussen groepen; voor en na de observatieperiode.

Contrair aan de verwachting wordt gevonden dat de leerlingresultaten over de gemeten periode dalen. Geen effect van de training op cognitieve en affectieve prestaties van leerlingen kon worden aangetoond. Wel wordt geconcludeerd dat de training geleid heeft tot verandering in leerkrachtgedrag (echter niet in overeenstemming met het ideale 'genormeerde' model).

Wanneer trainingscondities niet geheel beantwoorden aan verlangde kenmerken en het beheersingsgedrag van de leerkrachten niet voldoet aan de gestelde norm, is het moeilijk uit te maken waar (eventueel gevonden) verandering in leerprestaties aan kan worden toegeschreven. De auteurs nemen enkele theoretische verklaringen door.

Het boek behandelt twee interessante onderzoeken en een belangwekkend thema voor onderzoek. De auteurs geven een beschrijving die dicht bij de

feitelijke onderzoeksuitvoering blijft. Dit is vooral interessant voor degenen die willen leren hoe dergelijk onderzoek wordt uitgevoerd en welke stappen of beslissingen daarbij aan de orde zijn. De evocatieve titel 'changing teaching for better learning' laat onverlet dat het weinig biedt voor degene die leerprestaties wil verbeteren.

H. H. Tillema

Literatuur

Fenstermacher, G. D., On learning to teach effectively from research on teacher effectiveness. In: A. Denham & C. D. Lieberman, *Time to learn*. Washington: NIE, 1981.

J. C. Sturm, *Een goede gereformeerde opvoeding*. Over neo-calvinistische moraalpedagogiek (1880-1950), met speciale aandacht voor de nieuw-gereformeerde jeugdorganisaties. J. H. Kok, Kampen, 1988, 283 pp., *f* 40,50, ISBN 90 242 5430 2.

Twee jaar geleden promoveerde Sturm aan de VU op het hier te bespreken onderzoek (promotores: Spiecker en Kuitert). Het gedistiancieerd en af en toe licht ironisch geschreven boek laat zich goed lezen. Zélf gelooft de auteur dat zijn vader, 'een mild mens...die tijdens zijn leven buitengewoon veel inzicht gekregen (had) in de calvinistische psyche' zijn boek 'met plezier' gelezen zou hebben (p. 5). Sturm bedoelde natuurlijk niet dat hij zelf zijn boek stilistisch geslaagd vindt, doch dat de inhoud ervan zijn vaders belangstelling zou hebben gehad. Ik zeg direct aan het begin echter maar vast dat in ieder geval ook de wijze waarop hij van zijn onderzoek verslag doet charme bezit. Hij gebruikt in zijn normale betoog bijvoorbeeld – en dan altijd onverwacht – woorden die ontleend zijn aan het 'gereformeerde taalspel'. Hij geeft in 'inheemse' termen samenvattingen. Zinnen als: 'De gereformeerde kerkelijke geschiedenis heeft...heel wat ervaringen met...meningsverschillen,...gewetensonderzoek, harte- en nierproeverij, tuchtoefening, afsnijding, uitstoting, scheuring en afscheiding' vertonen dezelfde couleur locale als teksten van renegaten-literaten (Wolkers, 't Hart, Siebeling).

Maar: ook over de *inhoud* van het boek kan men tevreden zijn. Sturm doet in zes hoofdstukken (die bij nader inzien beter als drie delen zijn te beschouwen) verslag van respectievelijk: 1. de geschiedenis van het calvinisme en van de nieuw-gereformeerde beweging daarbinnen (hoofdstukken 1 t/m 4); 2. de nieuw-gereformeerde 'theorie' achter de morele opvoeding (hoofdstuk 5); en 3. de praktijk van de nieuw-gereformeerde jeugdorganisaties (hoofdstuk 6). Een leessuggestie op p. 14 zegt dat de eerste vier hoofdstukken tezamen en elk der volgende los van elkaar toegankelijk zijn – wat inderdaad klopt.

De waarde van het boek ligt vooral in de heldere weergave van een complexe materie. De omvang van de bronnen is indrukwekkend (de bibliografie, onderverdeeld in naslagwerken, bronnen en secondaire literatuur, kent 30 dicht bedrukte pagina's). Het gaat om als: wetenschappelijke teksten en notulen van jongelingenverenigingen, brieven en ingezonden stukken, krantenberichten en verenigingsperiodieken, jubileumboeken, enzovoort. De auteur heeft dit alles goed in de greep en weet plausibel te maken dat hij de kern van het onderzoek kan beperken tot een adequaat afgegrensde historie (van 1880 - 1950).

Ten behoeve van zijn onderzoek operationaliseert hij een aantal begrippen. Zo onderscheidt hij tussen 'gereformeerd' en 'nieuw-gereformeerd' en tussen 'Gereformeerd' en 'gereformeerd' (p. 33). De term 'Gereformeerd' staat voor alles wat binnen het institutionele kader van 'nieuw-gereformeerd' valt (pp. 34 en 179), terwijl 'nieuw-gereformeerd' staat voor de bijzondere (grotendeel kuyperiaanse) ideologie binnen het 'gereformeerde' (annex het calvinisme) in het algemeen. Daarnaast is ook de term 'oud-gereformeerd' produktief (als aanduiding voor een meer bevindelijke variant binnen het calvinisme; p. 38). (Tot mijn verrassing – ik ben nu eenmaal, om het sfeervol te zeggen, een 'ongodist' – noteert Sturm 'God', doch verwijst naar Hem als 'hij', 'hem', 'zijn'. Omdat hij kennelijk zorgvuldig over het gebruik van boven- en onderkast nadenkt, zit ik hier dus met een vraag...)

Hij slaagt er in het ethos van de nieuw-gereformeerde beweging treffend te schetsen en weet aannemelijk te maken hoe zij zich van proto-zuil tot zuil ontwikkelt. Sturm meent dat de beweging in de 20ste eeuw 'versteent', wat hij tragisch vindt. Het begon immers 'allemaal als een poging om orthodoxe calvinisten te behoeden voor het isolement in een achterlijk godsdienstig getto. Het eindigde met het isolement in een veelvoudig verzuild Nederland' (p. 27/28). Gezien de periode die Sturm onderzoekt, lijkt het begrijpelijk dat hij de ontwikkeling tragisch vindt. Van verstening is inderdaad sprake, tenminste als we niet verder kijken dan 1950. De geschiedenis is dan echter niet afgelopen. Secularisatie en het samengaan van godsdienstig-politieke 'bloedgroepen' volgen bijvoorbeeld op de 'verstening'. Misschien is een en ander tegen de achtergrond van de huidige situatie minder dramatisch te beoordelen en heeft ze zelfs een positieve functie gehad. Wat Sturm daarvan zou kunnen denken, laat zich echter slechts raden.

Sturm onthult dat de nieuw-gereformeerde opvattingen over 'souvereiniteit in eigen kring' veel te maken hebben met waardering voor het *organische* karakter van samenlevingsvormen. Daarbij wijst hij er terecht op dat deze denktrant niet verwant is met

het organicistische denken binnen de Duitse idealistische filosofie.
Soms blijven constateringen al te zeer op zichzelf staan terwijl ze, in verband met elkaar gebracht, een interessant probleem zouden kunnen oproepen. Zo wijst Sturm hiér op de souvereiniteit in eigen kring en het organische karakter van die kringen (samenlevingsvormen); zo wijst hij dáár op de beduchtheid voor staatsinvloed en op het voortdurende pleidooi voor vrijheid (met name vrijheid ván staatsinvloed: burgers moeten hun eigen idealen kunnen vormgeven; p. 42); en zo noemt hij élders de behoefte om de ganse samenleving te 'herkerstenen' (p. 67). Zo'n ('zo een', zou Sturm schrijven) herkerstende samenleving (Kuypers ideaal) is echter, consequent geredeneerd, een organische, op basis van één geloofsbelijdenis functionerende samenleving. Binnen die samenleving kan maar één vrijheid heersen: de 'vrijheid voor het juiste handelen naar Gods geboden' (p. 42). Mijns inziens moet dit betekenen: weg democratische vrijheid!
Deze frictie tussen de vrijheidsideologie en de 'herkersteningsideologie' signaleert hij niet. Hij wijst zonder commentaar op het neo-calvinistische democratiseringsethos en de zojuist genoemde, *wel heel specifieke* vrijheidsopvatting (p. 42). De (voor een historicus interessante) vraag naar de verhouding tussen deze singuliere vrijheidsoptiek ('vrijheid voor het juiste handelen naar Gods geboden') enerzijds en de praktisch-politieke effecten ervan in een situatie waarin een machtige calvinistische minderheid in een samenleving de politieke en bestuurlijke toon zet anderzijds, blijft eveneens onder tafel. (Denk aan de Duitsland- en economische politiek van de regeringen-Colijn, de eind-jaren-30-houding tegenover uit Duitsland vluchtende joden, de 'Indië-kwestie', enzovoort). Ook de historisch en analytisch interessante vraag naar de verhouding tussen burgerlijke vrijheid en gehoorzaamheid (vergelijk de 'Schakel- vs. Aantjes-houdingen'), die nauw samenhangt met het vorige probleem, blijft achterwege (al hoort dit thema strikt genomen misschien niet meer tot zijn onderzoek). De vraag naar de logische positie van het vrijheidsbegrip binnen het geheel der gereformeerde ideologie wordt door Sturm wél bekeken, maar alleen op het moment dat de verhouding tussen predestinatie en vrijheid besproken wordt.
Deze en andere punten van kritiek raken kwesties die vallen binnen de grenzen van het onderwerp en van de methodologische aanpak daarvan. Het gaat niet om principiële omissies; ze vallen op als je Sturm loyaal volgt op zijn onderzoeksweg. Men kan echter ook de gekozen onderzoeksmethoden als het ware van buitenáf beoordelen, bijvoorbeeld om te kijken of de gevolgde weg de meest produktieve is. Doe je dat, dan krijg je van lieverlee een probleem in het vizier (dat bij nader inzien reeds verborgen is in de compositie van het boek). Los van elkaar worden immers de *specifieke geschiedenis van het calvinisme (inclusief de nieuw-gereformeerde beweging)* (deel 1, bestaande uit 4 hoofdstukken), de *theoretische achtergronden* (deel 2 = hoofdstuk 5) en de *idee en praktijk van het derde milieu* (deel 3 = hoofdstuk 6) gegeven. Dat doet vermoeden dat de ene thematiek niet of nauwelijks met de twee andere in verband wordt gebracht. En dat ís ook zo.
Aanvankelijk had ik dit niet door. Wel stelde ik op p. 63 in de kantlijn de vraag waarom Sturm zegt Hendriks niet te zullen volgen. Op dat moment lanceert hij namelijk zijn plan om een overzicht te geven van de nieuw-gereformeerde geschiedenis, en wel *aan de hand van het sociaal-historisch perspectief* dat Hendriks in zijn dissertatie *De emancipatie van de gereformeerden* (1971) hanteert. Zo'n sociaal-historische kenschets géeft Sturm ook; het maakt hoofdstuk 4 (van 'deel 1') uit. Tegelijkertijd echter zegt hij Hendriks' sociaal-historische benadering niet te kunnen volgen bij de schets van de ontwikkeling van de nieuw-gereformeerde pedagogiek in theorie en praktijk (van de jeugdbeweging), te geven in de volgende twee grote hoofdstukken ('delen'). Hij beperkt zich dan ook inderdaad tot uiteenzettingen van de nieuw-gereformeerde ideologie 'als één geheel rond een aantal sleutelbegrippen' (p. 63). In feite licht Sturm hier dus een tipje van de sluier op die lag over de reeds op p. 14 van zijn boek gegeven leeswijzer: men kan, zegt hij daar, de hoofdstukken 1/4 en hoofdstuk 5 en hoofdstuk 6 als afzonderlijke delen lezen. Nú, op p. 63, begrijpt men waarom dat kan.
De geschiedschrijving, de theorie-analyse, en de beschrijving van de jeugdbewegingspraktijk staan dus ieder apart. 'Toch is het van belang voor een goed begrip van de nieuw-gereformeerde opvoeding', aldus Sturm, 'om kennis te hebben van het bredere historische kader' dat hij in deel 1 geeft. Deze toevoeging krijgt men er gratis bij. Het goede begrip moet de lezer namelijk zelf maken: de theoretische achtergronden en de inhoud van de nieuw-gereformeerde ideologie en praktijk (de omvangrijke hoofdstukken 5 en 6) worden niet in een groter historisch verband geplaatst. Sturm beschrijft de ideologie en, dat moet zonder meer erkend worden, doet dat helder en overtuigend. Ook aan de traditionele theoretische lijnen, vanaf Calvijn, wordt aandacht geschonken. Maar in hoeverre de ideologie en de daarop gebaseerde praktijk aan de ene kant een *reactie* is op maatschappelijke ontwikkelingen, en aan de andere kant *effecten* sorteert in de vorm van gereformeerd-politieke en -maatschappelijke posities en activiteiten, blijft onderbelicht.
Is, nog dieper gegraven, Sturms keuze voor een *driedelig* in plaats van een op historisch en theoretisch niveau *samenhangend* betoog nader te verklaren? Ik vermoed van wel, en meen te begrijpen dat het probleem ligt bij de strikt theoretische omschrijving van 'morele opvoeding' die hij als analytisch zoeklicht hanteert. Reeds op p. 9 definieert hij morele opvoeding als: 'de opvoeding om het goede te kennen, te kunnen doen en te wille kiezen'. Later geeft hij weliswaar uitgebreidere omschrijvingen,

maar in de kern gaat het om wat hij hier zo concies formuleert. Deze definitie is uiterst formeel. Met zo'n 'operator' in de hand laten zich de morele dimensies van élke ideologische pedagogiek herkennen, opgeacht de historische context. Die context komt namelijk, zoekend naar de invullingen van 'het goede', de manieren waarop je dit goede kunt 'kennen' en de wijzen waarop opvoeders bij hun kinderen aan 'wil'svorming werken, nooit in het vizier.

De gebruikte omschrijving bezit nog een ander euvel. Ook *als theoretische definitie* sec heeft ze namelijk haar beperkingen. Met behulp van een dergelijke formele omschrijving van morele opvoeding is elke (*externe*) kritiek onmogelijk. Of het nu om de pedagogiek en de opvoeding van de Komsomoljeugd, de Hitlerjugend, de nieuw-gereformeerde jeugd, de jeugd uit de 'zwarte kousen-kerken' gaat: alles wat daar 'goed', 'kennis' en 'willenswaard' heet, past er in. Als men dan alles wat op basis van de definitie tot morele opvoeding te rekenen is, verzameld heeft, kan men hooguit nog *immanente kritiek* leveren (bijvoorbeeld als de begripsmatige consistentie van een morele opvoedingsideologie niet deugt omdat concepten 'intern' strijdig zijn).

Deze theoretische indispositie van Sturms omschrijving van morele opvoeding wordt nog versterkt omdat Sturm onder 'kennis' van alles vat. 'Kennisgebieden' binnen de gereformeerde morele opvoeding zijn volgens Sturm: '– geloofsbelijdenis, ... kennis van de ... nieuw-gereformeerde geloofsovertuigingen; – historische onderwerpen, verdeeld in gewijde geschiedenis (Bijbel) ...; – maatschappelijke onderwerpen (hoofdzakelijk beginselen en standpunten van de Anti-Revolutionaire Partij...); – ethiek (uitgaande van de Tien Geboden); – letterkunde (christelijk proza ...)' (p. 185). Ook elders in het boek wordt de term 'kennis' in deze ruime betekenis van 'weet hebben van' gebruikt. Hoewel Sturm wel een enkele keer wijst op het indoctrinatiekarakter van de gereformeerde opvoeding – en dus tussen 'kennen' en 'geïndoctrineerd zijn' onderscheidt – ontbreekt hier elke analyse. (Hoe staat het met de geldigheid van wat men gelooft, welke metafysische en andere assumpties spelen daarbij een rol, enzovoort.) Dat is jammer, omdat een precies en kritisch gebruik van 'kennis' meer mogelijkheden biedt voor een beoordeling van wat in morele zin 'goed' en 'willenswaard' genoemd wordt in een ideologie.

Al met al lijkt het er op dat het gebruik van de definitie van morele opvoeding gevolgen heeft gehad voor zowel het theoretische als het historische gehalte van het onderzoek. Ik haast me echter te zeggen dat mijn 'externe' kritiek vooral als voer voor een nadere gedachtenvorming rond methodologische kwesties bij cultuurhistorisch onderzoek moet worden beschouwd. Per slot van rekening gaat het bij Sturms studie om een heldere presentatie van een hoogst interessant en typisch Nederlands verschijnsel. Dat 'om' zijn onderzoek 'heen' nog belangrijke vragen te stellen zijn, is eigenlijk een extra verdienste van het boek.

J. D. Imelman

T. Baudoin, T. Meelis-Voorma, P. Moolenaar, H. Overmeijer & B. Ros (red.) *Jeugdliteratuur voor de beroepspraktijk. Een leergang voor het HBO*. Wolters-Noordhoff, Groningen, 1990, 185 pag., *f* 39,95, ISBN 90 01 09223 3.

In 1989 werd "Adviezen over de voorlopige eindtermen basisonderwijs en basisvorming in het voortgezet onderwijs" door het Ministerie van Onderwijs en Wetenschappen aangeboden. Deze eindtermen waren al direct aanleiding tot heftige discussies. Allerlei disciplines zochten naar vooral hun eigen plaats binnen de eindtermen. Niet verwonderlijk dat de werkgroep 'Kinder- en jeugdliteratuur in het HBO' het rapport intensief bestudeerde. Al jaren waren er signalen dat de aandacht voor kinder- en jeugdliteratuur afnam op de opleidingen voor onderwijsgevenden. Opmerkelijk dat in dit advies jeugdliteratuur expliciet genoemd wordt en daardoor de status heeft gekregen die het volgens de werkgroep al jaren verdiende.

Met "Jeugdliteratuur voor de beroepspraktijk' presenteren Baudoin e.a. een zeer praktisch gerichte leergang voor het HBO. Het boek valt in drie delen uiteen. In het eerste deel 'Verhaal en lezer' geeft Moolenaar een korte schets van de geschiedenis van de jeugdliteratuur. Zoals de auteurs in de inleiding opmerken, ligt de nadruk op 'kort', want het belang van kennis van de geschiedenis wordt niet altijd onderkend. Moolenaar deelt de geschiedenis in vier periodes in: tot 1750, 1750-1800, 1800-1900 en 1900-1945. De periode na 1945 komt volgens de auteurs voldoende in de andere hoofdstukken aan de orde. Van elke periode worden achtereenvolgens behandeld: 'hoe zagen opvoeders het kind?', 'hadden kinderen een eigen leescultuur?', 'welke boeken lazen kinderen?' en 'welke eisen stelde men aan de lectuur voor kinderen?'. Alles wordt zeer bondig, puntsgewijs weergegeven. Elke periode wordt afgesloten met een werkopdracht voor de studenten. Deze bestaat uit het vergelijken van oude kinderboeken met moderne kinderboeken of uit het maken van een advieslijst voor het negentiendeeeuwse klassieke boek.

Eveneens in deel 1 een hoofdstuk over 'de lezer' en over 'beoordeling en selectie' (Ros). In deze hoofdstukken komt het leesgedrag, de invloed van de omgeving hierop, de leesmotivatie en de leesvoorkeuren van kinderen aan de orde. De reacties van de lezer tijdens het lezen worden teruggevoerd

op het centrale begrip 'identificatie'. Deze kan plaatsvinden op grond van medelijden, herkenning en gelijkenis en/of bewondering. Over de invloed van het gelezene is Ros korter. Hierover is ook veel minder bekend.

Dat beoordeling en selectie nauw samenhangen wordt duidelijk aangegeven. De traditionele indeling met betrekking tot het beoordelen wordt gehanteerd: de lezersgerichte, de maatschappijgerichte en de boekgerichte benadering. E.e.a wordt aan de hand van voorbeelden duidelijk gemaakt. Het hoofdstuk wordt afgesloten met een overzicht van beoordelingsschema's uit de vakliteratuur.

Was deel 1 vooral theoretisch van aard, deel 2 'Verhaal en boek' is praktisch van aard en zal voor de beoogde doelgroep aanleiding kunnen zijn tot uitwerking van verschillende werkopdrachten. Dit deel is bijna geheel van de hand van Meelis-Voorma. Zij begint met een uiteenzetting van uitgangspunten en doelstellingen en baseert zich daarvoor op vooral discussienota's inzake het leesonderwijs. Daarna volgt een aantal hoofdstukken waarin werkvormen worden aangeboden om jeugdboeken in het onderwijs te gebruiken. Zij start met de, helaas te vaak vergeten, werkvorm: vertellen en voorlezen. Hoe kies je een verhaal?, hoe vertel je een verhaal?, hoe pas je een verhaal in in je onderwijsprogramma?, hoe weet je of kinderen een verhaal leuk vinden?, slechts enkele van de steeds zeer praktische vragen die Meelis aan de orde stelt. Bewust kiest zij bij de behandeling van deze vragen voor een onderscheid tussen vertellen en voorlezen. Aansluitend wordt in hoofdstuk 6 een aantal literaire genres besproken, voorafgegaan door een bespreking van het begrip 'fantasie'. Ook prentenboeken komen uitgebreid aan de orde. Het laatste hoofdstuk van deel 2 gaat over 'het peilen van de leesbelangstelling' (Baudoin). Methoden om achter de leesbereidheid, de leesmotivatie en de leesvoorkeur te komen worden zeer kort besproken.

Deel 3 heeft de vorm van een naslagwerk(je). Allereerst wordt een thematisch geordend literatuuroverzicht gegeven, gevolgd door instellingen in Nederland en Vlaanderen die zich bezighouden met kinder- en jeugdliteratuur. Het boek wordt besloten met een hoofdstuk 'tekstrepertoire'. Een opsomming van kinder- en jeugdboeken, naar leeftijd geordend, waarvan de auteurs vinden dat minstens de helft in ieder geval door de studenten gelezen moet worden. Het betreft het 'goede' kinder- en jeugdboek. Aangezien kinderen (zeer) vaak ook grijpen naar triviale boeken, wordt hier kort aandacht aan besteed.

"Jeugdliteratuur voor de beroepspraktijk" is een zeer bruikbaar boek voor de verschillende HBO-opleidingen. Over de opzet is goed nagedacht. Het feit dat het boek voorbereid (en geschreven) is door een werkgroep van auteurs die normaal verschillende benaderingswijzen hanteren, heeft geresulteerd in een studieboek dat evenwichtig van opzet is. Zonder te kiezen voor een bepaalde manier om kinder- en jeugdboeken te beoordelen, worden diverse schema's naast elkaar gezet. In de verwerkingsopdrachten moeten de studenten zich deze schema's eigen maken en tenslotte kiezen voor een eigen schema (opgebouwd uit elementen van de diverse benaderingen).

Theorie en praktijk staan in een juiste verhouding tot elkaar. De verwerkingsopdrachten, een belangrijk aspect van deze leergang, geven hieraan een extra dimensie. Bij elke opdracht staat aangegeven hoeveel tijd de opdracht studenten kost. Daarnaast is er een differentiatie aangebracht met betrekking tot deze opdrachten. Hierdoor is het mogelijk het boek op verschillende opleidingen te gebruiken, iedere opleiding kan die opdrachten kiezen die voor de toekomstige beroepspraktijk van belang zijn.

De literatuursuggesties (onder meer aan het eind van het boek) maken het mogelijk dat studenten bepaalde aspecten nader uitwerken. Ook het hoofdstuk over tekstrepertoire zal voor de docenten een belangrijk hulpmiddel zijn. Niet alleen om een literatuurlijst voor de studenten samen te stellen, maar misschien ook om de bibliotheek van de opleiding van meer boeken te voorzien. Het overzicht van instellingen en recensenten had van mij achterwege mogen blijven. Dergelijke overzichten veranderen elke dag; ik ontdekte nu al verschillende fouten.

In het hoofdstuk over de geschiedenis van het kinder- en jeugdboek zijn pagina 27 en 28 verwisseld. Verder heeft het onderwijskundige voordelen om puntsgewijs de boven aangehaalde vragen te beantwoorden. Het gevaar van een dergelijke benadering ligt bij de docenten die dit boek zullen gaan gebruiken. Voor een tentamen levert dit hoofdstuk eenvoudig na te kijken vragen.

'Jeugdliteratuur voor de beroepspraktijk' zal voor de opleidingen een zeer praktisch hulpmiddel zijn. Voor zowel docenten als studenten biedt het voldoende handvatten om de jeugdliteratuur die plaats te bieden binnen de toekomstige onderwijspraktijk die zij verdient.

A. W. M. Duijx

M. K. Singley & J. R. Anderson, *The transfer of cognitive skill*. Harvard University Press, Cambridge, Massachusetts, 1989, VIII + 300 pag., $ 35.95, ISBN 0 674 90340 4 (via Bookimpex nr. 754773/307009).

Singley en Anderson passen in hun boek de ACT*-theory (spreek uit: 'act-star', het letterwoord wordt niet verklaard) over het verwerven van cognitieve vaardigheden toe op het transferprobleem: het toe-

passen van reeds verworven vaardigheden/kennis in nieuwe situaties. De auteurs onderscheiden bron- en doelkennis/-vaardigheden: de bronkennis/-vaardigheden worden naar de doelkennis/-vaardigheden getransfereerd. Beide soorten kennis/vaardigheden kunnen declaratief (op een of andere wijze cognitief gerepresenteerd) of procedureel zijn. Deze procedurele vaardigheden zijn hoofdzakelijk van de soort: 'als x en y en z, doe dan a of b' etc. Zij noemen een dergelijke procedureregel een produktie, zodra de genoemde variabelen een concrete waarde aannemen.

Het eerste hoofdstuk plaatst het transferonderzoek in historisch perspectief. De transfertheorie van Thorndike wordt besproken, evenals de opvattingen over transfer van Piaget, van de Gestalt-psychologen, van Osgood en van Gagné. Op basis van de lessen die uit deze bespreking getrokken worden presenteren de auteurs hun ACT*-theory-toepassing, door hen genoemd de 'Identical Productions Theory of Transfer'. Stelde Thorndike aan het begin van deze eeuw dat transfer wordt bevorderd door gelijke stimulus-respons combinaties in de bron- en doelsituatie, dan stellen Singley & Anderson nu dat transfer wordt bevorderd door identieke produkties in beide situaties.

De hoofdstukken 2 t/m 7 bevatten alle beschrijvingen van onderdelen van deze theorie en, vooral, beschrijvingen van de vele experimenten die gedaan zijn om de theorie te toetsen. Aan de orde komen: laterale transfer, negatieve transfer, de gebruiksspecificiteit van procedurekennis, het stimuleren van analoge transfer en declaratieve transfer.

In het dan volgende hoofdstuk (8) wordt, op basis van de uitkomsten van de experimenten met betrekking tot de genoemde onderwerpen, de balans opgemaakt voor de identieke produktietheorie.

In een slothoofdstuk (9) wordt nog ingegaan op de belangrijke rol van de (declaratieve) cognitieve representaties bij de aanvang van het transferproces.

Het aantal in deze publikatie behandelde onderwerpen noch het aantal besproken experimenten laat het geven van een (korte) samenvatting van het boek toe: daarvoor zijn het er (veel) te veel. Ik wil daarom volstaan met de vermelding van de belangrijkste bevindingen van de auteurs.

Gegeven de brontaak en de doeltaak waarin transfer van de bronkennis/-vaardigheden zou moeten optreden, is er aanvankelijk sprake van een positieve laterale transfer (verticale transfer komt nauwelijks aan de orde), in zoverre de taken een gemeenschappelijke declaratieve basis bezitten. In een vervolgfase is positieve transfer vooral afhankelijk van de gemeenschappelijkheid in produkties die nodig zijn om de taken te uitvoeren. De declaratieve kennis is in het algemeen vrij toegankelijk voor de leerder; een kernprobleem voor de leerder is echter, wáár hij in zijn declaratieve geheugen moet zoeken naar de voor de doeltaak relevante kennis. Als belangrijkste zoekstrategie wordt de analogie genoemd (ongeveer zoals in 'exemplarisch leren'). Veelal wordt de relevante analogie door de leerders niet spontaan gezien en is hulp in de vorm van aanwijzingen noodzakelijk. Is de relevante analogie tussen bron- en doeltaak evenwel gevonden in het declaratieve geheugen, dan kan op basis daarvan overgegaan worden tot het formuleren van doelsituatie-specifieke produktieregels, waarna de produkties die deze regels beschrijven worden uitgevoerd.

Om tot deze bevindingen te komen hebben de auteurs een groot aantal experimenten uitgevoerd; deze experimenten hadden betrekking op de transfer bij het zich eigen maken van verschillende tekstverwerkingsprogramma's, eenvoudige computertalen, differentiaal en integraal rekenen, en logisch en mathematisch probleem-oplossen.

Singley en Anderson behandelen in hun boek een zeer belangrijk onderwerp: de transfer van kennis en vaardigheden naar taken in doel- en toepassingssituaties. Dat zijn de situaties waarvoor uiteindelijk alle onderwijs bedoeld is. Opvallend is dat aan een zo essentieel onderwerp tot heden toch betrekkelijk weinig aandacht besteed is. Alleen daarom al is het boek van Singley en Anderson van belang. Bewondering verdienen de ingenieuze experimenten waarmee zij hebben geprobeerd het optreden van transfer te onderzoeken. Daaruit blijkt bovendien hoe moeilijk dit onderwerp zich voor onderzoek leent en hoeveel assumpties gemaakt moeten worden om het überhaupt te kunnen onderzoeken. Een meer nauwkeurige definitie van de onderzochte variabelen en hun operationalisaties zou, mede gelet op de vele assumpties, de onderzoekers gesierd hebben. Verschillen tussen kernvariabelen als kennis, vaardigheden, cognitieve vaardigheden, cognitieve representaties e.d. blijven nu nogal eens in het vage en daardoor kunnen hun operationalisaties moeilijk beoordeeld worden (de auteurs stellen daartoe ook geen pogingen in het werk).

De flaptekst belooft ons dat de bevindingen uit dit boek belangrijke implicaties voor de verbetering van het onderwijs zullen hebben. Deze implicaties, waarin ik zeer geïnteresseerd ben, heb ik helaas niet kunnen ontdekken. Dat is op zichzelf begrijpelijk, gezien de prille fase waarin het onderzoek en de theorievorming zich nog bevinden. Effectieve toepassing ervan op korte termijn in de onderwijspraktijk lijkt me niet goed mogelijk. Ondanks deze bedenkingen: een interessant boek over een boeiend en belangrijk onderwerp.

J. H. Slavenburg

Mededelingen

Landelijke Pedagogendag 1991: oproep tot het houden van voordrachten

De vijfde landelijke pedagogendag voor theoretisch en historisch opvoedkundigen wordt gehouden op zaterdag, 25 mei 1991 aan de Universiteit van Amsterdam. Onderzoek(st)ers presenteren in voordrachten van maximaal 20 minuten hun onderzoek, toegespitst op het (ruim op te vatten) thema 'Individuatie en socialisatie in tijden van modernisering'.

Voordrachthoud(st)ers moeten zich vóór 1 januari 1990 opgeven bij 'Commissie Pedagogendag, t.a.v. Frieda Heyting, vakgroep Pedagogische Wetenschappen U.v.A., IJsbaanpad 9, 1076 CV Amsterdam, met vermelding van de titel van de te houden voordracht. Samenvattingen voor het programmaboekje van maximaal 350 woorden (meer worden niet afgedrukt) op flop (liefst WP) moeten vóór 1 februari worden gestuurd naar hetzelfde adres.

Inhoud andere tijdschriften

Comenius
10e jaargang, nr. 3, 1990

Volwassenen-educatie in de jaren negentig, door F. Meijers
De Katholieke Film Centrale en haar keuringsbeleid tussen 1945 en 1967, door E. Radius
Staat en kind in Australië, 1840-1890, door R. van Krieken
Over wetenschapsgeschiedenis. Naar aanleiding van de lezing 'Beginsel en beroep' van Ernst Mulder, door M. Depaepe
Discussie: Het 'Onderwijs en Sociaal Milieu'-project mislukt? door W. Meijnen
Buitenlandse bladen
Boekbesprekingen

Pedagogisch Tijdschrift
15e jaargang, nr. 4, 1990

De implementatie van curriculum-materiaal voor maatschappijleer, door W. Kuiper en J. van den Akker
De balans als diagnosticum. Een onderwijsexperiment, door M. Wolters en J. Zuidema
Schoolmoeheid in het secundair beroepsonderwijs. De visie van de leerkracht, door D. Berings en F. de Fruyt
Besprekingsartikel: Spelling en spellingproblemen, door J. Rispens
Kroniek: Balans voor de toekomst. Impressies van het 12e congres van de Deutsche Gesellschaft für Erziehungswissenschaft te Bielefeld, 19-21 maart 1990, door S. Miedema en G. J. J. Biesta
Boekbesprekingen

Pedagogisch Tijdschrift
15e jaargang, nr. 5, 1990

Computersimulatie in de cognitieve psychologie, door K. Lamberts
De techniek van computersimulatie in het onderzoek naar het leren oplossen van vraagstukken over optellen en aftrekken, door L. Verschaffel en E. De Corte
Functies en structuur van 'Intelligent Tutoring Systems' door J. Breuker, R. Winkels en A. Muntjewerff
Naar een flexibele methode voor algoritmisch grammatica- en spellingonderwijs, door G. Kempen en E. Jongen-Janner
Mens en machine. Anti-mechanistische argumenten in de filosofische discussie rond Artificiële Intelligentie, door H. Roelants
De computer als simulator en als tutor van onderwijsleerprocessen, door L. de Klerk en L. Verschaffel.

Tijdschrift voor Onderwijsresearch
15e jaargang, nr. 4, 1990

Effecten van een training in basale gespreksvaardigheden, door S. M. Bögels en R. Kreutzkamp
Gebruik van verzorgingsactiviteiten in het basisonderwijs, door D. Koster, Th. van Batenburg en H. H. Tillema
De betrokkenheid van schoolleiders in het V.O. bij vernieuwingen in het onderwijs, door P. C. van de Sijde en E. Pruis
Notities en Commentaren
Onderwijsonderzoek te kijk gezet. Commentaar op het rapport 'Schaal en kwaliteit in het basisonderwijs' door B. M. P. Creemers en W. Th. J. G. Hoeben
Boekbesprekingen

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 9, 1990

Incorrecte afspraken over individueel diagnostisch onderzoek ten behoeve van justitiële instanties, door M. K. van der Heijden
Diagnostiek, hulpverlening en justitie, door H. Baartman
Het 'agogisch' kennisbelang als inzet, door W. van der Schee
Onderscheid tussen aandacht- en leesproblemen bij kinderen, door E. A. Das-Smaal, E. F. J. M. Brand en S. L. van den Hooff
De valkuil van het didactische leeftijdsequivalent, door J. H. L. Oud en M. J. C. Mommers

Tijdschrift voor Orthopedagogiek
29e jaargang, nr. 10, 1990

Themanummer interculturele pedagogiek, door L. Eldering
De Marokkaanse en Nederlandse gezinscultuur: relationele en rationele stimulatie, door T. Pels
Hulpverleningsdiagnositek bij jeugdigen uit allochtone groepen, door P. J. A. Adriani
Interventieprogramma's voor jonge kinderen uit etnische minderheidsgroepen, door L. Eldering

Ontvangen boeken

Abma, R., *Jeugd en tegencultuur. Een theoretische verkenning* (dissertatie). Uitgeverij SUN, Nijmegen, 1990, *f* 29,50.

Beijaard, D., *Teaching as acting* (dissertatie). Vakgroep Agrarische Onderwijskunde, Landbouwuniversiteit Wageningen, 1990, *f* 30,–.

Berg, R. van den (red), *Bouwstenen voor evaluatie voor de praktijk van onderwijs en begeleiding.* Bohn Stafleu Van Loghum, Houten, 1990, *f* 45,–.

Boei, F., *Ontwikkelingen in declaratieve kennis* (dissertatie). Academisch Boeken Centrum, De Lier, 1990.

Driel, J. van, *Betrokken bij evenwicht* (dissertatie). Centrum voor Didactiek van Wiskunde en Natuurwetenschappen, R.U. Utrecht, 1990, *f* 20,–.

Jansen, A. W. H., *'Ik zal je mores leren!' Een onderzoek naar de relatie tussen opvoedergedrag en de morele ontwikkeling van het kind* (dissertatie). Katholieke Universiteit Nijmegen, 1990

Kemme, S., *Uitleggen van wiskunde* (dissertatie). Vakgroep OW & OC, Rijksuniversiteit Utrecht, 1990.

Liere, C. van, *Lastige leerlingen. Een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen.* Uitgeverij Thesis, Amsterdam, 1990, *f* 35,–.

Pieters, J. M., P. R. J. Simons, L. de Leeuw (Eds.), *Research on computer-based instruction.* Swets & Zeitlinger, Lisse, 1990, *f* 48,50.

Thoomes, D. Th., *Opgroeien in perspectief. Theologisch-pedogogische opstellen.* Uitgeverij J. H. Kok, Kampen, 1990, *f* 14,90.

Tillekens, G. (red.), *Nuchterheid en nozems. De opkomst van de jeugdcultuur in de jaren vijftig.* Uitgeverij Coutinho, Muiderberg, 1990, *f* 34,50.

Siebenheller, E., *Problematische opvoedingssituaties: percepties, emoties en disciplineringsreacties van ouders* (dissertatie) 1990.

Snik, G., *Persoonwording en opvoeding* (dissertatie). Mondiss Uitgeverij, Kampen, 1990, *f* 45,–.

Veerman, J. W., *De ontwikkeling van kinderen na een periode van klinische jeugdhulpverlening* (dissertatie). Acco, Amersfoort, 1990, *f* 35,–.

Verwaaijen, A. A. G., *Therapist behavior in the treatment of families with adolescent girls at risk of placement* (dissertatie). Katholieke Universiteit Nijmegen, 1990.

Weenink, J. B. (red.), *Het belang van het kind in het geding. Over opvoeding en kinderbescherming.* VU Uitgeverij, Amsterdam, 1990, *f* 19,50.

Zande, I. van der, *De jongerenwerker als cultureel erflater. Het praktisch en theoretisch leerrendement van een handelingsonderzoek* (dissertatie). Acco, Amersfoort, 1990, *f* 29,–.

In 1990 werd redactionele medewerking verleend door:

E. M. H. Assink (R.U. Utrecht)
Th. Bergen (K.U. Nijmegen)
R. J. Bosker (U.T. Enschede)
M. Depaepe (K.U. Leuven)
J. Dronkers (U.v.A. Amsterdam)
T. J. Ferguson (Utah State University, Logan)
W. van de Grift (R.U. Utrecht)
A. W. van Haaften (K.U. Leuven)
J. D. Imelman (R.U. Utrecht)
R. de Jong (RION, Groningen)
P. Jungbluth (ITS, Nijmegen)
F. van der Krogt (K.U. NijmegenJ. M. G. Leune (E.U. Rotterdam)
H. van der Meij (U.T. Enschede)
W. A. J. Meijer (R.U. Utrecht)
J. W. Mojet (R.U. Leiden)
M. J. C. Mommers (K.U. Nijmegen)
C. A. N. van Opdorp (K.U. Nijmegen)
H. M. Pijnenburg (Paedologisch Instituut, Nijmegen)
G. C. de Vries (S.C.O., Amsterdam)
O. Wiegmans (U.T. Enschede)
J. M. Wijnstra (CITO, Arnhem)
Th. Wubbels (R.U. Utrecht)