



Pedagogische studiën 1996 : Tijdschrift voor onderwijskunde en opvoedkunde

<https://hdl.handle.net/1874/205428>

Pedagogische Studiën

Tijdschrift voor onderwijskunde en opvoedkunde

Inhoud Jaargang 73 1996

Bibliotheek der
Rijksuniversiteit
UTRECHT

RIJKSUNIVERSITEIT TE UTRECHT



2737 520 7

Pedagogische Studiën

Tijdschrift voor onderwijskunde
en opvoedkunde

Redactie

Prof. Dr. N. Verloop (voorzitter)

Dr. P. J. J. Stijnen (secretaris)

Dr. K. P. van den Bos

Prof. Dr. J. J. Hox

Prof. Dr. W. Koops

Prof. Dr. G. W. Meijnen

Prof. Dr. P. R. J. Simons

Dr. J. N. Streumer

Prof. Dr. L. Verschaffel

Dr. J. M. Wijnstra

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

Uitgever: Wolters-Noordhoff, Postbus 567, 9700 AN Groningen, tel. 050-522 68 86
Correspondentie over abonnementen en losse nummers uitsluitend aan de uitgever.

Inhoudsopgave

Jaargang 73, 1996

Auteursregister

Artikelen

Aarts, R., Ruiter, J.J. de, & Verhoeven, L.
Eigen-taalvaardigheid en schoolsucces van
Turkse en Marokkaanse leerlingen 275

Appel, R., & Vermeer, A.
Uitbreiding van de Nederlandse woorden-
schat van allochtone leerlingen in het basis-
onderwijs 82

Bakkenes, I.
zie: Imants, J.

Berg, R. van den
zie: Geijssel, F.

Blok, H., & Leseman, P.P.M.
Effecten van voorschoolse stimuleringspro-
gramma's: een review van reviews 184

Blom, S.
zie: Wardekker, W.

Booij, N.
zie: Houtveen, A.A.M.

Brabander, C. de
zie: Imants, J.

Busato, V.V.
zie: Prins, F.J.

Corte, E. De
zie: Verschaffel, L.

Felling, A.J.A.
zie: Slik, F.W.P. van der

Geijssel, F., Berg, R. van den, & Slegers, P.
Het innovatief vermogen van scholen in het
basisonderwijs: een tweede vooronder-
zoek 42

Gerris, J.R.M.
zie: Slik, F.W.P. van der

Grift, W.J.C.M. van de
zie: Houtveen, A.A.M.

Hamaker, C.
zie: Prins, F.J.

Hofman, R.H.
Effecten van het verschil in bestuursvorm
voor het openbaar en bijzonder basisonder-
wijs 93

Hofman, R.H., & Meijer, C.J.W.
Weer Samen Naar School: uitkomsten van
onderzoek en toekomstperspectief. Intro-
ductie op het thema 402

Hofman, R.H., & Vonkeman, E.B.
Onderwijsmethoden en adaptief onder-
wijs 411

*Houtveen, A.A.M., Booij, N., Jong, R. de, &
Grift, W.J.C.M. van de*
Adaptief onderwijs en leerlingresulta-
ten 422

Imants, J., Brabander, C. de, & Bakkenes, I.
Isolement en taakperceptie van schoollei-
ders in basisscholen 8

Jong, R. de
zie: Houtveen, A.A.M.

Jong, S. de, & Riemersma, A.M.J.
Taalpeiling in Friesland 291

Jonghe, H. de
zie: Veenman, S.

Kinket, B.
zie: Verkuyten, M.

Klatter, E.B.
Studievaardigheden in de brugklas. Een
vergelijking tussen algemene en vakspeci-
fieke verwerkingsstijlen van brugklasleer-
lingen 303

- Krüger, M.L.*
Sekse van de schooldirecteur in relatie tot de gerichtheid op het primaire proces 25
- Leseman, P.P.M., & Şahin, S.*
Tweetalige ontwikkeling van Turkse peuters 259
- Leseman, P.P.M.*
zie: Blok, H.
- Meijer, C.J.W.*
Weer Samen Naar School in cijfers 404
- Weer Samen Naar School: terugblik en vooruitzicht 441
- Meijer, C.J.W.*
zie: Hofman, R.H.
- Oudenhoven, D.*
zie: Petersen, B.
- Petersen, B., & Oudenhoven, D.*
Waar komen regionale verschillen in verwijzing vandaan? 434
- Prins, F.J., Busato, V.V., Hamaker, C., & Visser, K.H.*
Een bijdrage tot de validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen 108
- Raemaekers, J.*
zie: Veenman, S.
- Riemersma, A.M.J.*
zie: Jong, S. de
- Ruiter, J.J. de*
zie: Aarts, R.
- Şahin, S.*
zie: Leseman, P.P.M.
- Sleegers, P.*
zie: Geijsel, F.
- Slik F.W.P. van der, Gerris, J.R.M., & Felling, A.J.A.*
Opleiding van ouders en de opvoeding van hun kinderen. Een onderzoek in de Kohn-traditie 338
- Steutel, J.W.*
De deugdenbenadering van de morele opvoeding. Een poging tot conceptuele verheldering 215
- Stijnen, P.J.J.*
zie: Verhoeven, L.
- Veenman, S., Jonghe, H. de, & Wezel, E. van*
Evaluatie van een trainingsprogramma voor het coachen van leraren 170
- Veenman, S., & Raemaekers, J.*
Retentie-effecten van een nascholingsprogramma voor effectieve instructie en klasmanagement 357
- Verhoeven, L., & Stijnen, P.J.J.*
Aspecten van tweetaligheid. Introductie op het thema 250
- Verhoeven, L.*
zie: Aarts, R.
- Verkuyten, M., Kinket, B., & Wielen, Ch. van der*
Oudere kinderen en hun begrip van discriminatie 447
- Vermeer, A.*
zie: Appel, R.
- Verschaffel, L., & Corte, E. De*
Leren realistisch modelleren en interpreteren van vraagstukken. Een exploratief onderwijsexperiment bij leerlingen van de bovenbouw van de basisschool 322
- Visser, K.H.*
zie: Prins, F.J.
- Vollebergh, W.A.M.*
Autoritarisme en etnocentrisme bij leerlingen uit lagere opleidingsniveaus: instemmingstendens of opinie? 123

Discussiebijdragen

Vonkeman, E.B.

zie: Hofman, R.H.

Voort, T.H.A. van der

zie: Walma van der Molen, J.H.

Walma van der Molen, J.H., & Voort, T.H.A. van der

Leereffecten van Jeugdjournaal en gedrukt nieuws bij kinderen: Een mediumvergelijkend experiment 459

Wardekker, W., & Blom, S.

In memoriam Prof.Dr. J.F. Vos 6

Wezel, E. van

zie: Veenman, S.

Wielen, Ch. van der

zie: Verkuyten, M.

Wieringen, A.M.L. van

In memoriam Prof.Dr. Ph.J. Idenburg 2

Witziers, B.

Coördinatie van het vakonderwijs binnen scholen voor voortgezet onderwijs: vaksecties en schoolleiding 198

Gemert, G.H. van

Prikkels tot zorgzaam onderwijs? 143

Kroonenberg, P.M., & Veer, R. van der

Internationale publikaties en kwaliteit. Een onderzoek naar het publikatiegedrag van Nederlandse pedagogen en onderwijskundigen 225

Meijer, C.J.W.

Realisering Weer Samen Naar School-beleid afhankelijk van wijziging bekostigingssysteem 136

Reactie op de bijdragen van A.J.J.M. Ruijsenaars en G.H. van Gemert 146

Redactie Pedagogische Studiën

De onderzoeksvisitatie 1995 239

Ruijsenaars, A.J.J.M.

Weer Samen Naar School: als het doel op is, beginnen we aan de middelen 140

Veer, R. van der

zie: Kroonenberg, P.M.

Kronieken

Knoers, A.M.P.
Onderzoek beoordeeld 56

Pieters, J.M. (red.)
Zesde EARLI-Conferentie. Nijmegen,
26-31 augustus 1995, met bijdragen van: *J.*
Andriessen, J. van der Linden, J.M. Pieters,
G. Seegers, H.H. Tillema, M. Valcke, M.V.S.
Veenman 60

Streumer, J.M. (red.)
Annual Meeting American Educational Re-
search Association 18-22 april 1995 San
Francisco, met bijdragen van: *M. Mulder,*
J.M. Pieters, M.G. Roes, J.N. Streumer, H.J.
Vos, H. de Vos 149

Verloop, N., & Verschaffel, L. (red.)
Het AERA-Congres 1996, New York, met
bijdragen van: *D. Beijaard, E. De Corte,*
B.P.M. Creemers, F. Dochy, J. Elen,
J.F.M.J. van Hout, E. Knoors, J. Lowyck,
G.W. Meijnen, M. Mulder, N. Verloop, J.
Vermunt, L. Verschaffel 372

Verschaffel, L.
zie: Verloop, N.

Boekbesprekingen

Eerde, H.A.A. van
Kwantiwijzer. Diagnostiek in reken-
wiskundeonderwijs (*J.M.C. Nelissen*) 474

Geus, W.C. de
Op zoek naar functionele leesvaardigheid
(*H.C. Lutje Spelberg*) 161

Heuvel-Panhuizen, M. van den
Assessment and realistic mathematics edu-
cation (*J.M.C. Nelissen*) 474

Hox, J.J.
Applied Multilevel Analysis (*T. Mooij*)
..... 244

Keulen, H. van
Making sense: Simulation-of-Research in
organic chemistry education (*J.H. van*
Driel) 75

Lowyck, J., & Verloop, N. (Red.)
Onderwijskunde: een kennisbasis voor pro-
fessionals (*G. Kanselaar en E. Roelofs*)
..... 158

Peschar, J., & Wesselingh, A.
Onderwijssociologie (*S. Karsten*) 242

Rijt, A.M. van
Voorbereidende rekenvaardigheid bij kleu-
ters. De ontwikkeling van rekenvaardig-
heidsschalen en een onderzoek naar de in-
vloed van een programma (*J.M.C. Nelissen*)
..... 474

Theunissen, M.W.G.
Samen nascholen, maar hoe? (*J. Imants*)
..... 317

Thema: Schoolleiderschap

Nummer 1

- J. Imants, C. de Brabander en I. Bakkenes*
Isolement en taakperceptie van schoolleiders in basisscholen 8

- M.L. Krüger*
Sekse van de schooldirecteur in relatie tot de gerichtheid op het primaire proces 25

Thema: Aspecten van tweetaligheid

Nummer 4

- L. Verhoeven en P.J.J. Stijnen*
Aspecten van tweetaligheid: Introductie op het thema 250

- P.P.M. Leseman en S. Şahin*
Tweetalige ontwikkeling van Turkse peuters 259

- R. Aarts, J.J. de Ruiter en L. Verhoeven*
Eigen-taalvaardigheid en schoolsucces van Turkse en Marokkaanse leerlingen 275

- S. de Jong en A.M.J. Riemersma*
Taalpeiling in Friesland 291

Thema: Weer Samen Naar School

Nummer 6

- R.H. Hofman en C.J.W. Meijer*
Weer Samen Naar School: uitkomsten van onderzoek en toekomstperspectief. Introductie op het thema 402

- C.J.W. Meijer*
Weer Samen Naar School in cijfers 404

- R.H. Hofman en E.B. Vonkeman*
Onderwijsmethoden en adaptief onderwijs 411

- A.A.M. Houtveen, N. Booij, R. de Jong en W.J.C.M. van de Grift*
Adaptief onderwijs en leerlingresultaten 422

- B. Petersen en D. Oudenhoven*
Waar komen de regionale verschillen in verwijzing vandaan? 434

- C.J.W. Meijer*
Weer Samen Naar School: terugblik en vooruitzicht 441

5 oct 6441

4B-lu L

Inhoudsopgave

Pedagogische Studiën Jaargang 73, nummer 1/1996

A. M. L. van Wieringen

In memoriam Prof. Dr. Ph. J. Idenburg 2



W. Wardekker en S. Blom

In memoriam Prof. Dr. J. F. Vos 6

J. Imants, C. de Brabander en I. Bakkenes

Isolement en taakperceptie van schoolleiders in basisscholen 8

M. L. Krüger

Sekse van de schooldirecteur in relatie tot de gerichtheid op het primaire proces 25

F. Geijssel, R. van den Berg en P. Sleegers

Het innovatief vermogen van scholen in het basisonderwijs: een tweede vooronderzoek 42

Kronieken

A. M. P. Knoers

Onderzoek beoordeeld 56

J. M. Pieters (red.)

Zesde EARLI-Conferentie. Nijmegen, 26-31 augustus 1995 60

Boekbespreking 75

Mededelingen 78

Maatschappijwetenschap en onderwijs

Eind vorig jaar overleed op 94 jarige leeftijd Philip J. (Philippus Jacobus) Idenburg. Het onderwijs en de statistiek in Nederland zijn hem veel verschuldigd. Idenburg begon zijn loopbaan als adjunct-secretaris van de Schoolraad van de Scholen met de Bijbel in 1924. Hij leerde daar aan den lijve de wetstoepassing op basis van de onderwijspacificatie kennen. Onderwijsbeleid was in die tijd beperkt tot deze toepassing. Hij zou dat later aanduiden als de periode van allocatief onderwijsbeleid en pleiten voor een constructiever onderwijsbeleid.

Hij promoveerde in 1928 op een proefschrift over de staat en het volksonderwijs. In 1929 begon hij bij het Centraal Bureau voor de Statistiek als hoofd van een nieuwe afdeling onderwijsstatistiek. Hij bleef daar afgezien van een kleine maar wel belangrijke onderbreking werkzaam tot 1966. Die onderbreking betreft de functie van directeur-generaal voor het onderwijs in 1946 onder de eerste socialistische minister van onderwijs, Van der Leeuw. De KVP minister Gielen die Van der Leeuw opvolgde gaf Idenburg zijn congé. Als belangrijkste opbrengst van die korte periode beschouwde hijzelf de benoeming van Max Goote tot inspecteur-generaal. Max Goote zou later bij de totstandkoming van de wet op het voortgezet onderwijs een essentiële rol vervullen. Hij keerde weer terug naar het CBS waar hij reeds in 1939 was benoemd tot directeur-generaal van de statistiek. Mede via de opbouw van de onderwijsstatistiek heeft hij historische, sociologische en ook economische benaderingen van het onderwijs tot ontwikkeling gebracht.

Aanvankelijk met enige tegenzin werkzaam op het ogenschijnlijk saaie gebied van het onderwijs kreeg hij gaandeweg meer en meer belangstelling voor het terrein. Ook de activiteiten van zijn schoonvader Philip Kohnstamm, hoogleraar in de pedagogiek aan de Universiteit van Amsterdam, waren behulpzaam bij het ontstaan van enthousiasme voor het onderwijs. Idenburg is van 1934 tot 1972 met de Universiteit van Amsterdam verbonden geweest. In 1934 werd Idenburg als privaat-docent in de geschiedenis, theorie en statistiek van het

schoolwezen toegelaten aan de Universiteit van Amsterdam. In 1956 werd hij bijzonder hoogleraar tevens directeur van het Nutsseminarium. Deze directeursfunctie werd min of meer volgens gebruik gecombineerd met het voorzitterschap van de redactie van Pedagogische Studiën. Een functie die hij tot 1972 vervulde. Onder zijn leiding veranderde het blad ook officieel van een pedagogisch orgaan tot een tijdschrift voor onderwijskunde en opvoedkunde. In 1965 werd hij de eerste voorzitter van de Stichting voor Onderzoek van het Onderwijs. Een Stichting die veel heeft betekend voor de opbouw en de verbetering van de onderwijsresearch in Nederland. Na zijn pensionering als directeur-generaal was hij van 1967 tot 1972 gewoon hoogleraar in de maatschappelijke achtergronden van opvoeding en onderwijs en de vergelijkende opvoedkunde (later omgezet in: algemene en vergelijkende onderwijskunde).

Idenburg is veel geweest, maar ook een aantal dingen niet: geen minister van onderwijs (ondanks de bedoeling van Vorrink) en geen burgemeester van Amsterdam (de ministerraad gaf de voorkeur aan de andere kandidaat: Van Hall). Hij heeft het niet betreurd. Waarschijnlijk ook omdat het met Van Hall niet echt goed afliep.

In wetenschappelijk opzicht is Idenburg vooral bekend vanwege een tweetal boeken. In de eerste plaats de *Schets van het Nederlandse schoolwezen* uit 1960, in 1964 geheel herzien. In de wandeling nog altijd aangeduid als *De Schets*. In de tweede plaats is dat *Theorie van het Onderwijsbeleid* uit 1971. In *De Schets* geeft hij een analyse van de factoren die ons onderwijs vormgeven zoals centralisatie, verscheidenheid en ongelijkheid. Hij benadrukte dat het onderwijs als één geheel moest worden bestudeerd en bestuurd: destijds een noviteit. In die periode werd in het voortgezet onderwijs aan aantal schoolsoorten in één wettelijk kader geregeld. Een dergelijke onderneming werd toen zo spectaculair bevonden dat men er de naam mammoetwet aan gaf. Beide boeken ge-

ven een brede maatschappij-analyse van het onderwijs waarin historische, sociologische, politicologische en economische benaderingen geïntegreerd worden gehanteerd. In de onderwijsgeschiedenis, de onderwijssociologie en in de studie van het onderwijsbeleid zijn deze boeken nog steeds van belang. Minder is dat het geval in de onderwijsconomie. Toch is bijvoorbeeld zijn inaugurale rede uit 1967 *Het schoolwezen als economisch verschijnsel* nog steeds een fraai verwoorde problematisering van deze benadering. Idenburg was een maatschappijwetenschapper die de verschillende disciplines zoals sociologie, politicologie en economie geïntegreerd behandelde. Hij vertegenwoordigde het ideaal van een geïntegreerde maatschappijwetenschap. Economische bijdragen zijn gedurende zijn gehele loopbaan verweven met sociologische en politicologische bijdragen. Hij schreef reeds in 1930 over de economie van het onderwijs in zijn *Is ons onderwijs de kosten waard?* In dat artikel stelt Idenburg: 'De economische vraagstelling is niet allesbeheersend en zij mag het nimmer worden. Maar, dat de vraag naar kost en baat van ons schoolwezen in onzen tijd haar recht heeft, ja meer dan tot dusver de ernstige aandacht van ons volk verdient, daarvan hoop ik den lezer te hebben overtuigd'. In 1940 publiceerde hij over *School en efficiëncy* en in 1961 schreef hij over *Onderwijs en welvaart*. Bijna veertig jaar na 1930 komt hij in zijn hierboven reeds vermelde aanvaardingsrede van 1967 *Het schoolwezen als economisch verschijnsel* opnieuw te spreken over de economie van het onderwijs vooral aan de hand van de opkomende Chicago school met mensen als Becker, Denison, Schultz en Bowman. Er zijn weinig mensen in Nederland die deze brede behandeling destijds op waarde schatten.

Hij schreef in een ouderwets mooi Nederlands voor zijn blad *Pedagogische Studiën* veel artikelen en nog meer boekbesprekingen. Veel van zijn lezingen en wetenschappelijke bijdragen zijn in dit tijdschrift verschenen. Bijdragen over historische onderwerpen zoals *Thorbecke's middelbaar onderwijswet* (1963), over beroeps pedagogische onderwerpen zoals *Technisch onderwijs in een veranderende maatschappij* (1960) en over de onderwijsstatistiek in *Dynamische onderwijsstatistiek*

(1967). Afleveringen waarin hij drie à vier boeken besprak waren geen uitzondering. In de besprekingen ging hij verder dan signaleren en beoordelen. Sommige publikaties werden ook besproken in het licht van gewenste veranderingen in het beleid. Het jaarlijkse Onderwijsverslag had zijn aandacht. 'In een document van zo hoge standing mogen wij een geborneerdheid als de gesignaleerd niet aantreffen. Wij zouden anders werkelijk nog gaan geloven dat het ministerie wat de studie van de onderwijsproblematiek aangaat wat 'onderontwikkeld' is en dat het de zaken bovendien te veel van één kant bekijkt', schreef hij over het Onderwijsverslag van 1959.

Zijn artikelen gingen veelal over de onaangepastheid van het onderwijs aan de maatschappelijke ontwikkelingen en over de overheid die hier veel te laks optrad. Ik noem onder meer zijn aanvaardingsrede uit 1956 getiteld *Mensen gevraagd!* en vlak daarop volgend in 1958 zijn bijdrage met de veel geciteerde term *De sleutelmacht der school* (1958). In 1962 behandelde hij nogmaals wat het centrale thema zou worden van de Nederlandse onderwijssociologie: *Het ideaal van de optimale ontwikkeling der talenten en de pedagogische structuur van het onderwijs*.

Idenburg heeft zich ontwikkeld vanuit een christelijk-historische achtergrond naar een sociaal-democratische gezindheid. Hij werd na de oorlog lid van de Partij van de Arbeid. Deze overgang hing samen met zijn visie op het onderwijs en vooral op een actiever onderwijsbeleid. Nadat de KVP minister Gielen hem terugstuurde naar het Centraal Bureau voor de Statistiek bleef hij, vanuit met name zijn posities aan de Universiteit van Amsterdam, het onderwijsbeleid volgen. Hij ergerde zich aan het amateurisme waarmee onderwijs werd bestuurd door vooral traditioneel werkende uitvoerende onderwijsjuristen in plaats van door meer toekomstgerichte wetenschappelijk geschoolde ambtenaren. Dat amateurisme klemde temeer omdat de sociale en economische betekenis van het onderwijs in de decennia na de oorlog sterk was toegenomen. Het zou tot aan het begin van de jaren zeventig duren voor dat zijn analyses en pleidooien in daden werden omgezet.

Als het onderwijs zo iets kent als een geheugen dan is Idenburg in de herinnering het meest bekend vanwege zijn formulering van het begrip constructieve onderwijspolitiek. In zijn artikel *Naar een constructieve onderwijspolitiek* (1970) werd het duidelijkst kritiek geleverd op het onderwijsbeleid. Teveel werd het beleid ingeperkt tot het alloceren van geld aan scholen en te weinig werden problemen als ongelijkheid en vernieuwing aangepakt. Hij pleitte ook in zijn *Theorie van het onderwijsbeleid* voor een constructief onderwijsbeleid in plaats van een allocatief onderwijsbeleid. 'In de constructieve onderwijspolitiek wordt de overheid bij haar bemoeiingen met het schoolwezen door uitgesproken doeleinden geleid en de activiteit tot innovatie van het onderwijs in al zijn geleidingen neemt in de strategie een wezenlijke plaats in' (Idenburg, 1971, p. 167).

Dit pleidooi had succes omdat het op een gedegen analyse van de tekortkomingen van het beleid was gebaseerd en omdat de tijd überhaupt rijp was voor een actiever overheidsbeleid. Vooral PvdA minister Van Kemenade werd de vertolker van een dergelijke constructieve onderwijspolitiek. Hij beoogde het schoolwezen uitdrukkelijker in te zetten om maatschappelijke doelen als gelijkheid en weerbaarheid te bevorderen. Het achterstandenbeleid, de middenschool en de open school zijn bijvoorbeeld in het kader van die constructieve politiek ontworpen. De term constructief onderwijsbeleid is tijden lang synoniem geweest met goed onderwijsbeleid. Zelfs de opvolger van minister Van Kemenade de VVD minister Pais kon er niet goed onderuit, het door hem in gang gezette emancipatiebeleid was bijvoorbeeld uitdrukkelijk constructief in de Idenburgse opvatting.

Idenburg formuleerde een opvatting over hoe het onderwijsbeleid gevoerd behoorde te worden die vooral door minister van Kemenade is gepraktiseerd. Van Kemenade en Idenburg hebben meerdere dingen gemeen. Beiden zijn doorbraakachtige politici resp. bestuurders. Van Kemenade is overigens niet alleen in politiek opzicht onderwezen in de opvattingen van Idenburg. Na zijn ministerschap was hij onder meer hoogleraar onderwijskunde en in die hoedanigheid zorgde hij in 1981 voor de opvolger van Idenburgs Schets met het boek

getiteld *Onderwijs: Bestel en Beleid*.

Idenburg was als hoogleraar een beminnelijk mens met veel gevoel voor understatement. Samen met Bas van Eijndhoven (de latere directeur-generaal voor het basisonderwijs) vormde ik de categorie aangeduid als 'mijn beide assistenten'. Het leek hem nuttig als we ook zijn colleges bijwoonden en dus zaten Bas van Eijndhoven en ik wekelijks op de eerste rij. Meestal in gezelschap van een paar studenten die te laat binnen waren gekomen en door Idenburg steevast met veel vriendelijkheid waren uitgenodigd om als tegenprestatie voorin te komen zitten. Het hielp wel. Wij staken inderdaad wat op van de colleges en het aantal te laat komende studenten nam af.

De betekenis van Idenburg ligt in het snijvlak van analyse en handeling. Zijn wetenschappelijke arbeid was niet zozeer gericht op verklaring maar vooral op verbetering van het handelen. Het door hem geformuleerde begrip constructief onderwijsbeleid bevatte beide aspecten: het was een analytisch begrip en tegelijkertijd een opvatting over hoe er gehandeld moest worden.

In zijn *Schets* en in zijn verschillende lezingen heeft hij gepleit voor wat tegenwoordig zou heten een systeemanalyse van het onderwijs gekoppeld aan een professionalisering van het onderwijsbeleid. De enorme expansie van het onderwijs en de groeiende maatschappelijke betekenis van het onderwijs stond in geen verhouding met de kortzichtige-traditionele beleidsvoering van het onderwijs. Onderwijs werd, in zijn woorden, beheerd als een snoepwinkel.

Zijn pleidooien voor een beter onderwijsbeleid gingen uit van zowel een behoefte aan een deskundiger vorm van onderwijsbeleid als een behoefte aan een socialer vorm van onderwijsbeleid. Lang niet al zijn tegenstanders konden het onderscheid tussen beide overwegingen aanbrengen. In zijn afscheidsrede *De utopie in het onderwijsbeleid* van 1972 formuleert hij nogmaals zijn positie: een omvattende maatschappijanalyse gericht op verbetering van het handelen. De rede handelt over de actualiteit van de utopie. De utopie kent radicale varianten. "Maar de kennis der historie en het begrip van mens en samenleving leren ons dat wij ge-

noegen moeten nemen met een aanval, die stap voor stap voorwaarts gaat.... Maar laat ons de utopie bewaren als draagster van ons diepste verlangen en als richtpunt van ons beleid! Zij is de drijfveer van alle vooruitgang en niet de geringste onder de uitingen van een doorleefde menselijkheid”.

Idenburg was een veelzijdig mens. Hij was geïnteresseerd in cultuur en kunst. Onderwijs was voor hem geen wereld op zich zelf. Onderwijs was onderdeel van een breder cultuurbegrip. Hij schreef over de *Taak en vorming van de intellectueel in het nieuwe Europa* (1961). Hij was ook actief op het terrein van de kunst. Zo was hij voorzitter van het Nederlands Film-instituut en van de Raad voor de Kunst.

In 1973 emigreerde hij naar Wallonië en deed rigoureuus afstand van al zijn vakboeken.

De tuin en de literatuur en vooral de vele bezoekers uit de kring van familie, vrienden en kennissen kregen nu de aandacht. Zijn mooi opgeknapte boerenhuis in de Ardennen zal velen nog goed voor ogen staan evenals de wandelingen door het heuvelachtige landschap met halverwege het onvermijdelijke aanbrennen van de reep chocolade. De laatste paar jaren woonde hij opnieuw in Wassenaar waar hij op 29 december 1995 overleed.

In 1976 heb ik aan zijn huisuitgever voorgesteld een aantal van zijn klassieke artikelen te bundelen. De uitgever ging om commerciële redenen niet in op het voorstel. Twintig jaar later hebben deze artikelen niet aan betekenis ingeboet en het voornemen om ze bijeen te brengen leeft nog steeds.

A. M. L. van Wieringen

Met Jaap Vos is, in alle stilte, een markant en veelkleurig onderwijswetenschapper heengegaan. Zijn werk en zijn persoon hadden verschillende, soms tegenstrijdige kanten. Hij was altijd op de hoogte van de laatste maatschappelijke en politieke trends, maar mocht zich ook graag omringen met grote stapels filosofische en historische boeken. Net als bij zijn leermeester Leon van Gelder getuigt zijn werk zowel van eruditie als van maatschappelijke bewogenheid. Hij had een voorkeur voor het grootse en omvattende. Tegelijk was hij wars van te ver doorgedreven idealisme, wat hem soms verleidde tot een eerder pragmatische houding.

Zijn dissertatie "Onderwijswetenschap en Marxisme" geeft blijk van zijn behoefte om ontwikkelingen in een breder kader te zien. Het was een van de eerste pogingen om het werk van Vygotskij en Leont'ev te plaatsen in de ontwikkelingen in de Sovjet-Unie rond de Russische Revolutie. Dat leidde tot een veel complexere, en vooral ook meer pedagogisch georiënteerde, interpretatie van wat bekend stond als de 'Russische leerpsychologie'. Met die interpretatie werd deze benadering van een technisch-didactisch hulpmiddel tot een alternatieve manier van denken over onderwijs die zich nog steeds verder ontwikkelt.

Ook later heeft hij in diverse publikaties aan een diepgaander begrip van die denkwijze bijgedragen, onder andere door de relaties bloot te leggen met de ontwikkeling van het pedagogisch denken in West-Europa, die voor hem haar expressie vond in het werk van Langeveld. Vos had daarbij al in zijn proefschrift duidelijk gemaakt dat hij van Marxisme of communisme als gesloten systemen niets moest hebben. Niet voor niets was hij gefascineerd door de kritisch pedagoog Heydorn, in wiens werk de afkeer van zulke systemen al centraal stond lang voordat dat 'postmodern' begon te heten. In deze scepsis paste ook kritiek op het in zijn ogen naïeve dualisme van de Reformpedagogische denkers. Hun vertrouwen in de 'natuurlijke' ontwikkeling van het kind die als vanzelf tot een betere wereld zou leiden, zoals Rousseau

dat al eerder had verwoord, droeg in zijn ogen het risico in zich van blindheid voor de perversitering van zulke utopieën.

In zijn theorievorming en onderzoekspraktijk rond de vormgeving van de Middenschool (ook een voortzetting van het werk van Van Gelder) trachtte hij beide aspecten, de Vygotskyaanse en pedagogische manier van denken over onderwijs en het wantrouwen tegen grote systemen, te concretiseren. Zijn Amsterdamse oratie en het rapport voor de WRR over basisvorming zijn daarvan voorbeelden, evenals de samenwerking met een nieuw opgerichte Middenschool te Almere. Als een der eersten benadrukte hij het belang van 'maximale, communale' eindtermen, waaruit het denken in termen van 'basisvorming' voortvloeide. In zijn werk bleek, naast zijn voorkeur voor grootse projecten en grote lijnen, ook duidelijk zijn voortdurende bewustzijn van de politieke haalbaarheid en (vooral) onhaalbaarheid van bepaalde ideeën. Hij koos dan een 'overstijgende benadering', die soms leidde tot vaagheden, compromissen en inconsistenties in zijn voorstellen. Maar consistent is wel de achterliggende gedachte die voortkomt uit maatschappelijke bewogenheid: de school moet werken aan de 'promotion de tous', alle leerlingen meer en betere mogelijkheden bieden.

Zo'n Middenschool moest in zijn optiek mede ontwikkeld worden door een interdisciplinaire onderwijskunde. Aan het opzetten daarvan heeft hij veel aandacht besteed. Het ging hem daarbij om een 'betere onderwijskunde, niet meer onderwijskunde', wat voor hem vooral betekende dat de meer technologisch georiënteerde manier van werken moest worden aangevuld of ingekaderd met een meer pedagogische visie waarin zowel maatschappelijke en historische kaders als het belang van het individuele kind aan de orde zijn. Meerdere malen heeft hij zijn twijfels uitgesproken bij de feitelijke ontwikkelingen in de onderwijskunde, die hem te veel gegrondvest waren in een positivistisch vooruitgangsgeloof en vertrouwen in de wetenschap. Dat deed hij bijvoorbeeld in het als basisboek bedoelde "Onderwijskunde, een

inleiding" dat hij samen met Lagerweij redigeerde.

Teleurgesteld was Vos wanneer niet iedereen de waarde van zijn ideeën direct inzag en klaarstond om er praktisch vorm aan te geven. De weerstand tegen zijn vaak grootschalige plannen en de twijfels rond het vage karakter van voorgestelde compromissen, waardoor er minder van de grond kwam dan hij zich voorstelde, verbitterden hem. Mede daardoor was het de laatste jaren stil geworden rond zijn persoon. Maar zijn werk is voor velen een bron van

inspiratie gebleven. Zeker de rond het werk van Vygotskij op gang gebrachte theorievorming staat, mede door parallelle internationale ontwikkelingen, volop in de belangstelling. En ook zijn bijdragen aan het denken over de vormgeving van het Nederlandse onderwijs laten hun sporen na.

Jaap Vos was van 1978 tot 1986 lid van de redactie van *Pedagogische Studiën*.

W. Wardekker
S. Blom

J. Imants, C. de Brabander & I. Bakkenes

Samenvatting

Isolement is een verschijnsel dat tot nu toe met name bij leerkrachten is onderkend en beschreven. Aan de centrale positie van schoolleiders binnen hun schoolteam wordt doorgaans grote waarde toegekend, met name met het oog op doelmatig onderwijskundig leiderschap. Ondanks dat heeft isolement bij schoolleiders tot nu toe nauwelijks in de belangstelling gestaan. Doel van het onderzoek waarover hier wordt gerapporteerd is om isolement van schoolleiders in basisscholen onderzoekbaar te maken, en om aanwijzingen op te sporen voor het taakspecifieke karakter daarvan. Isolement van schoolleiders wordt vastgesteld op basis van de positie van schoolleiders in de communicatienetwerken van hun scholen. Daarnaast wordt schoolleiders gevraagd naar hun taakperceptie. De vraag is of verschil in isolement van schoolleiders binnen hun schoolteam samenhang vertoont met hun perceptie van taken. De resultaten wijzen in de richting dat geïsoleerde schoolleiders taken op het terrein van onderwijskundig leiderschap minder positief waarnemen dan niet-geïsoleerde schoolleiders.

1 Inleiding

Isolement is een verschijnsel dat tot nu toe met name bij leerkrachten is onderkend en beschreven (Lortie, 1975; Rosenholtz, 1989). Bedoeld wordt een situatie waarin de leerkracht zich uitsluitend op het werk voor de eigen klas(-sen) concentreert of daarop teruggeworpen wordt, en nauwelijks tot geen professionele communicatie met collega-leerkrachten onderhoudt. Isolement van leerkrachten wordt in verband gebracht met verlies van motivatie voor het werk, vasthouden aan vertrouwde maar weinig effectieve routines en defensief gedrag ten opzichte van vernieuwingen (Lortie, o.c.; Ashton & Webb, 1986; Nias, Southworth & Yeomans, 1989; Rosenholtz, o.c.; Staessens, 1993).

Aan de centrale positie van schoolleiders binnen hun schoolteam wordt doorgaans grote waarde toegekend, met name met het oog op doelmatig onderwijskundig leiderschap. Het belang dat aan de centrale positie van de schoolleider in de school wordt toegekend garandeert echter allerm minst dat alle schoolleiders ook daadwerkelijk deze positie innemen (Fullan, 1991). Ondanks dat heeft isolement bij schoolleiders tot nu toe nauwelijks in de belangstelling gestaan. In analogie met ons onderzoek naar isolement bij leerkrachten (Bakkenes e.a., 1994; Imants & Bakkenes, 1993) gaan we er in dit artikel van uit dat isolement van de schoolleider tot uiting komt in een oriëntatie op een relatief smal pakket van taken die individueel kunnen worden uitgevoerd, en die niet uit hun eigen aard communicatie met leerkrachten inhouden; de directeur trekt zich als het ware terug in de eigen directiekamer. Schoolleiderstaken die op individualistische wijze kunnen worden uitgevoerd zijn beheersmatige en administratieve taken. Voor schoolleiders in basisscholen met lesgevende taken doet de mogelijkheid voor individualistische taakuitvoering zich eveneens voor bij taken gericht op het werk met de leerlingen. Daar tegenover staan taken op het terrein van onderwijskundig leiderschap, die een brede oriëntatie op het werk van schoolleider impliceren. Vaak zullen deze taken ook een zekere mate van communicatie met leden van het schoolteam met zich meebrengen. De veronderstelling die in dit artikel centraal staat is dat de mate van isolement van schoolleiders een samenhang vertoont met de wijze waarop de schoolleider omgaat met taken op het terrein van onderwijskundig leiderschap, en dat isolement van de schoolleider op gespannen voet staat met een doelmatige uitvoering van taken op het terrein van onderwijskundig leiderschap.

Doel van dit onderzoek is om isolement van schoolleiders in basisscholen onderzoekbaar te maken, en om aanwijzingen op te sporen voor het taakspecifieke karakter van isolement zoals

dat hiervoor kort is aangeduid. Isolement van schoolleiders wordt vastgesteld op basis van de positie van schoolleiders in de communicatienetwerken van hun scholen. Een lage mate van participatie in een netwerk wordt opgevat als een indicatie voor isolement. Taakperceptie wordt opgevat als een indicatie voor hoe schoolleiders met taken omgaan, meer in het bijzonder welke taken schoolleiders selecteren als zijnde taken behorend tot hun functie. Taakperceptie, taakselectie en participatie in een communicatienetwerk zijn begrippen die in de volgende paragraaf nader worden toegelicht. In het onderzoek waarover hier wordt gerapporteerd gaat het om de vraag of verschil in isolement van schoolleiders binnen hun schoolteam samenhang vertoont met hun perceptie van taken. De verwachting is dat geïsoleerde schoolleiders taken op het terrein van onderwijskundig leiderschap minder positief zullen waarnemen dan niet-geïsoleerde schoolleiders.

2 Isolement van directeurs in basisscholen

2.1 De context van isolement

Om de positie van schoolleiders binnen basisscholen te kunnen begrijpen is het nodig om enig oog te hebben voor het paradoxale karakter van de eisen die aan schoolleiders worden gesteld. Deels vloeien deze paradoxale eisen voort uit recente tendensen in het onderwijs, en deels liggen ze besloten in de context van het werk van schoolleiders.

Gedurende de afgelopen jaren is er sprake van een sterke profilering van de functie van schoolleider ten opzichte van leerkrachten. Een zekere afstand van de schoolleider ten opzichte van leerkrachten wordt hierdoor in de hand gewerkt. In de praktijk zal een aanzienlijk deel van de vergrote beleidsruimte voor scholen en schoolbesturen aan de schoolleider in mandaat gegeven worden. Op basis van dit zogenaamde directiestatuut zal de schoolleider tot op zekere hoogte de rol van werkgever gaan spelen ten opzichte van de leerkrachten. De schoolleider zal meer dan voorheen belast zijn met een groot aantal algemene en onderwijskundige managementtaken binnen en buiten de school. De functie van schoolleider is inhoudelijk, organisatorisch en rechtspositioneel sterker dan voorheen

onderscheiden van de functie van leerkracht (Imants, 1994). Dit maatschappelijk proces van profilering van het beroep van basisschool-directeur zal niet als vanzelf tot isolement van de schoolleider ten opzichte van leerkrachten leiden. Een zekere mate van afstand tussen directeur en leerkracht kan door dit proces echter wel in de hand worden gewerkt.

De schoolleider krijgt op basis van inzichten in het belang van onderwijskundig leiderschap ook te maken met de eis van betrokkenheid op het werk van de leerkracht. Zo krijgt de directeur van de basisschool door de Commissie Evaluatie Basisonderwijs (1994) een belangrijke rol toebedeeld bij het realiseren van een meer 'opbrengstgerichte cultuur' in basisscholen. De directeur wordt geacht bij het doorbreken van leerkracht-isolement en het ontwikkelen van zinvolle vormen van samenwerking een centrale rol te spelen (Firestone & Wilson, 1985). Profilering ten opzichte van leerkrachten en betrokkenheid op het werk van leerkrachten vertonen een zekere spanningsverhouding, waarbinnen de directeur van de basisschool zelf positie moet kiezen.

Op schoolleiders van basisscholen ligt een zware schoolexterne en -interne druk om zich met vele taken tegelijkertijd en naast elkaar bezig te houden. De gemiddelde schoolleider in het basisonderwijs werkt 50 uur per week aan de leidinggevende en lesgevende taken. De invulling van de functie met taken verschilt sterk van schoolleider tot schoolleider, en deze invulling vertoont geen direct verband met zaken als grootte van de school, aanwezigheid van een conciërge, etc. (Stoel, 1994). Autonomie en schaalvergroting van de school brengen een toenemend aantal beheersmatige en algemene managementtaken met zich mee, bijvoorbeeld op het terrein van personeelsbeleid en gebouwenbeheer. Volgens Stoel komen veel schoolleiders dan ook nauwelijks toe aan wat zij zien als hun primaire taak, namelijk de vormgeving van het onderwijskundig leiderschap van de school. De toenemende eisen aan zowel onderwijskundig leiderschap, als algemeen management plaatsen de directeur van de basisschool in een als tegenstrijdig ervaren situatie, waarin de schoolleider zelf positie moet kiezen.

Een oriëntatie op de eigen lesgevende taken en algemene managementtaken, in combinatie met een profilering van de functie, kan isole-

ment van de schoolleider ten opzichte van leerkrachten in de hand werken. Los van bovengenoemde ontwikkelingen ligt isolement deels in de functie van schoolleider besloten. Een school die doelmatig functioneert kan worden beschouwd als een school die in haar beleid en haar dagelijks functioneren een evenwicht vindt tussen de verwachtingen en behoeften van interne betrokkenen, zoals leerlingen, leerkrachten en schoolleiding, en externe betrokkenen, zoals ouders, overheid en instellingen voor vervolgonderwijs (Alblas & Van de Vliert, 1990). Het vinden en in de dagelijkse praktijk realiseren van een dergelijk evenwicht is een voor iedere school complexe en tot op zekere hoogte ook unieke opgave. Alle betrokken partijen zullen aan de schoolleider een belangrijke rol toekennen bij het ontwikkelen en instandhouden van dit evenwicht. Geen van de partijen zal zich daarbij echter als de natuurlijke bond- en deelgenoot van de schoolleider opwerpen (Fullan, 1991).

In het werk van de schoolleider zijn isolement en autonomie moeilijk te onderscheiden begrippen. In deze bijdrage gaan we ervan uit dat het als individu verantwoordelijkheid kunnen nemen voor allerlei beslissingen een voorwaarde is voor effectief leiderschap. Een schoolleider die al haar of zijn zorgen wil delen met collega-leerkrachten kan als een storende factor worden gezien voor het effectief functioneren van deze leerkrachten. De professionele autonomie van de schoolleider – het vermogen om als individu verantwoordelijkheid te nemen voor het eigen werk – kan als een positieve factor voor het doelmatig functioneren van de school worden beschouwd. Autonomie en isolement zijn termen die regelmatig door elkaar gebruikt worden, maar die verschillende betekenissen hebben. Mede met het oog op verheldering van deze concepten wordt in dit onderzoek getracht om isolement van schoolleiders in basisscholen meetbaar te maken. Daartoe wordt isolement opgevat als het resultaat van defensief organisatiegedrag.

2.2 Isolement en defensief organisatiegedrag

Flinders (1988) heeft voor de bestudering van isolement in scholen onder meer het ecologisch en het adaptieve perspectief onderscheiden. In het ecologisch perspectief is isolement inherent aan de algemene werkplaats- en beroeps-

condities in de school. Een voorbeeld van onderzoek dat vanuit dit perspectief is uitgevoerd is het eerder genoemde werk van Lortie, waaruit naar voren komt dat isolement van leerkrachten door algemene condities in het beroep en de school in de hand wordt gewerkt. Op vergelijkbare wijze is in de voorgaande paragraaf naar voren gebracht dat in het beroep van schoolleider in basisscholen tendensen tot isolement besloten liggen, die voor alle schoolleiders gelden. Onderzoek wijst echter uit dat tussen en binnen scholen leerkrachten en schoolleiders verschil vertonen in de mate van isolement (Bakkenes e.a., o.c.). Daaruit kan worden afgeleid dat condities in en rond het beroep en de school niet voor alle schoolleiders dezelfde gevolgen hebben, en dat schoolleiders daar op uiteenlopende wijze mee omgaan. Dit betekent dat het ecologisch perspectief op isolement niet zonder meer houdbaar is en nuancering behoeft.

In het adaptief perspectief op isolement wordt benadrukt dat isolement het gevolg is van defensief organisatiegedrag. Dit defensief gedrag is gericht op het veilig stellen van als schaars waargenomen en belangrijk geachte hulpbronnen voor het werk, die voor een adequate taakuitvoering noodzakelijk geacht worden (Ashforth & Lee, 1990). In het geval van leerkrachten kan bij deze schaars geachte hulpbronnen bijvoorbeeld worden gedacht aan tijd om ongestoord met leerlingen te kunnen werken. Defensief organisatiegedrag door de schoolleider wordt mede opgeroepen en mogelijk gemaakt door de tegenstrijdigheden in de hiervoor beschreven eisen die aan de schoolleider gesteld worden. Zoals eerder is aangeduid geven deze paradoxen ruimte voor een uiteenlopende invulling van de eigen functie. Wanneer de schoolleider voldoende is toegerust en een zekere controle heeft op de gegeven schoolcontext, kan hij of zij de geboden kansen in positieve zin aanwenden, en wordt de schoolleider de centrale spil in de school, ook in onderwijskundige zin. Daar staat tegenover dat schoolleiders in een minder gunstige startpositie zich primair zullen richten op die taken die in ieder geval moeten worden uitgevoerd om de school en de eigen klas draaiende te houden, te weten de lesgevende taken en de algemene managementtaken. Verwacht mag worden dat in deze situatie taken op het terrein van

onderwijskundig leiderschap, die veel interactie met leerkrachten impliceren, defensief gedrag in de hand zullen werken. Legitimering voor dit defensief gedrag is op voorhand aanwezig door de grote druk om zich met lesgevende en algemene managementtaken bezig te houden.

Aanwijzingen voor de houdbaarheid van deze zienswijze kunnen in enkele Amerikaanse onderzoeken worden gevonden naar de mate waarin leerkrachten betrokken worden bij het werk van de directeur. Deze resultaten bevestigen de zienswijze op isolement als resultaat van defensief organisatiegedrag; de schoolleider schijnt zich af van betrokkenheid van de leerkracht met het oog op de handhaving van diens gezag en status (Barth, 1987; Lindle & De Marco, 1988). Dit defensief gedrag is gericht op het beschermen van de eigen positie tegen de uithollende werking van de grote hoeveelheid uiteenlopende verwachtingen die binnen en buiten de school aan de directeur worden gesteld. Het resultaat van dit defensief gedrag is een zekere mate van isolement.

Het ecologische en het adaptieve perspectief staan ook voor verschillende zienswijzen op de wijze waarop schoolleiders omgaan met taken op het terrein van onderwijskundig leiderschap. Het ecologisch perspectief wordt verwoord door Stoel (o.c.), wanneer hij concludeert dat in het algemene schoolleiders vanwege condities in hun werk nauwelijks toekomen aan de vormgeving van onderwijskundig leiderschap. Vanuit het adaptief perspectief proberen we in dit artikel een verbinding te leggen tussen de mate van isolement van een schoolleider ten opzichte van leerkrachten en de wijze waarop schoolleiders onderwijskundige leiderschapstaken en andere taken waarnemen. Verondersteld wordt dat geïsoleerde schoolleiders hun werk dusdanig afgrenzen dat een relatief smal takenpakket ontstaat, waarin het accent in vergelijking met minder geïsoleerde collega's ligt op lesgevende en administratief/beheersmatige taken. Niet-geïsoleerde schoolleiders selecteren relatief meer taken die gericht zijn op sturing en verbetering van het onderwijsleerproces in de school, en worden gekenmerkt door een breed takenpakket. Met deze veronderstelling wordt een nuancering aangebracht in de conclusie van Stoel dat veel schoolleiders in basisscholen

de vormgeving van het onderwijskundig leiderschap zien als hun primaire taak. We willen niet bestrijden dat schoolleiders aan onderwijskundig leiderschap veel belang toekennen. We willen laten zien dat schoolleiders, in samenhang met hun positie binnen het schoolteam, verschil vertonen in de mate waarin ze gemotiveerd zijn om onderwijskundige leiderschapstaken als onderdeel van hun functie te selecteren.

2.3 Isolement taakspecifiek benaderd

Voor de taakspecifieke uitwerking van isolement van de schoolleider is het nodig om de begrippen taak, functie en taakdifferentiatie te omschrijven. Een taak is een afgerond geheel van arbeidsverrichtingen, bijvoorbeeld 'met een leerling zijn/haar ontwikkeling, prestaties en vorderingen bespreken', of 'het werven en selecteren van personeel'. Een functie, zoals de functie van directeur van een basisschool, is een verzameling van min of meer samenhangende taken. Daarbij kan een onderscheid worden gemaakt tussen een formele omschrijving van een functie en de feitelijke invulling van een functie door een individuele functionaris. Het is gebruikelijk dat bij een formele omschrijving van de functie van directeurs van basisscholen gestreefd wordt naar volledigheid bij het specificeren van de functie in onderscheiden taken (zie bijv.: Gemeente Rotterdam, 1990; De Wit, 1991). Anders dan door deze formele functie-omschrijvingen wordt gesuggereerd wijst onderzoek naar taakdifferentiatie in scholen uit dat binnen en tussen scholen grote verschillen bestaan in de feitelijke invulling van de functies van schoolleiders en leerkrachten met lesgevende en klasoverstijgende taken (I.T.S., 1975; Smets, 1983). Taakdifferentiatie is de verdeling van taken binnen scholen tussen individuen met verschillende of met identieke functies (Imants, 1986; 1994). Voortbouwend op het adaptief perspectief op isolement gaan we ervan uit dat schoolleiders, evenals leerkrachten, over een zekere speelruimte beschikken om taken te selecteren die onderdeel uitmaken van hun functie. Taakselectie wordt beschouwd als een individueel proces waarbij taken worden aanvaard als zijnde een onderdeel van het totale werk van de leerkracht of directeur (Imants & Bakkenes, o.c.). Aldus wordt elke schoolleider

gekaracteriseerd door een eigen configuratie van taken.

Voor de beschrijving en analyse van deze configuraties van taken wordt uitgegaan van het onderscheid binnen een schoolorganisatie in drie domeinen van besluitvorming en taakuitvoering: (1) het professionele domein van de leerkracht (en de schoolleider voor zover deze een lesgevende taak heeft), (2) het beheers- en bestuursdomein van de directeur, en (3) een gebied waar beide domeinen elkaar overlappen (Hanson, 1979; Imants, 1989; Schuit, Slegers & Giesbers, 1991). Taken op het gebied van onderwijskundig leiderschap bevinden zich met name in het gebied waar beide domeinen elkaar overlappen, omdat onderwijskundig leiderschap betrekking heeft op coördinatie en sturing van professionele werkzaamheden op klasoverstijgend en schoolniveau. Onderwijskundige leiderschapstaken kunnen zowel door schoolleiders, als door leerkrachten worden uitgevoerd, bijvoorbeeld in de functie van interne begeleider ('teacher leader') of senior-leerkracht.

We veronderstellen dat niet-geïsoleerde schoolleiders in vergelijking met geïsoleerde schoolleiders veel taken selecteren uit het gebied, waar de twee hiervoor genoemde domeinen elkaar overlappen. Een geïsoleerde directeur wordt gekenmerkt doordat deze zich in vergelijking met niet-geïsoleerde directeurs beperkt tot taken en beslissingen binnen het beheers- en bestuursdomein, aangevuld met taken uit het professionele domein voor zover het de lesgevende taak betreft. Niet-geïsoleerde schoolleiders worden gekenmerkt door een relatief breed takenpakket, waarin taken zijn opgenomen die gericht zijn op sturing en verbetering van het onderwijsleerproces in de school.

Motivatie voor taken wordt opgevat als een indicatie voor taakselectie. Deze motivatie wordt vastgesteld door schoolleiders te vragen naar de wijze waarop zij de betreffende taak waarnemen. Hiermee wordt voortgebouwd op het werk van Hackman en Lawler (1971) naar taakwaarneming en motivatie in werksituaties. Volgens deze auteurs is er sprake van een gunstige conditie voor motivatie wanneer een actor beschikt over het inzicht dat deze zelf verantwoordelijk is voor de uitvoering van taken die als betekenisvol worden gezien, en wanneer de actor in deze inzichten bevestigd wordt. De

wijze waarop een taak wordt waargenomen is een belangrijke conditie voor motivatie om deze taak uit te voeren (Bakkenes e.a., o.c.). Op grond van algemene inzichten in de samenhang tussen kenmerken van het werk en isolement (Daft & Macintosh, 1981; Dallinger, 1985, Nias, 1989; Rosenholtz, o.c.) worden in het onderzoek naar leerkracht- en schoolleidersisolement drie taakkenmerken centraal gesteld:

- 1 autonomie of zeggenschap: de mate waarin een werker zelf de werkwijzen, planning en doelen kan bepalen (eigen verantwoordelijkheid voor taak);
- 2 variatie: de variatie in activiteiten en hulpmiddelen bij het uitvoeren van werkzaamheden (betekenisvolheid van taak);
- 3 feedback op het eigen functioneren: de mate waarin de werker informatie ontvangt over de kwaliteit van het eigen functioneren (inzicht).

Mede op basis van de resultaten van onderzoek naar isolement en taakperceptie bij leerkrachten in basisscholen (Bakkenes e.a., o.c.) veronderstellen we dat de taakperceptie van schoolleiders voor lesgevende en beheersmatige taken weinig verschil vertoont, terwijl voor taken op het terrein van onderwijskundig leiderschap een samenhang gevonden wordt tussen mate van participatie door de schoolleider in het communicatienetwerk van de school en taakperceptie.

2.4 Communicatienetwerken van scholen

Centraal in professioneel isolement staat het achterwege blijven van taakgerichte communicatie met collega's. Taakgerichte communicatie heeft betrekking op onderwerpen die direct met het werk te maken hebben, en wordt onderscheiden van communicatie op sociaal-emotioneel terrein. Dit onderscheid is van belang, omdat een bepaald type scholen voorkomt waarin teamleden met elkaar een intensief sociaal-emotioneel contact hebben, zonder diepgaand met elkaar over het werk te praten. Staessens (1989) spreekt in dit verband van 'familiescholen'. In het onderzoek naar professioneel isolement in scholen is met name de taakgerichte communicatie van belang, omdat we taakgerichte communicatie beschouwen als een indicator voor isolement. Verschillen in taakgerichte communicatie tussen en binnen

scholen kunnen zichtbaar en meetbaar gemaakt worden met zogenaamde communicatienetwerken.

Een communicatienetwerk geeft het patroon van communicatie binnen een school weer; het vat het communicatiegedrag van alle teamleden binnen de school samen (Farace, Monge & Rusell, 1977). De positie van een individu binnen een communicatienetwerk wordt in essentie bepaald door de frequentie waarmee dit individu communiceert en het aantal collega's waarmee het individu communiceert. De resulterende netwerken verschaffen informatie op het niveau van de organisatie en het individu. Op het niveau van de organisatie kan bijvoorbeeld de dichtheid van het netwerk worden vastgesteld, waarbij scholen kunnen variëren in de mate van dichtheid. Een eilandschool zal een lage mate van dichtheid kennen, terwijl een collegiale school een hoge mate van dichtheid heeft.

Om de positie van een individu binnen een communicatienetwerk vast te stellen is in dit onderzoek de mate van participatie door schoolleiders in het netwerk van de school berekend. Een lage mate van participatie in het communicatienetwerk zien wij als een indicator voor isolement. Verschillen tussen schoolnetwerken en posities van schoolleiders daarbinnen worden in dit artikel geïllustreerd door drie netwerken met verschillende posities van directeuren te bespreken.

3 Methode

3.1 Opzet van het onderzoek

Voor het vaststellen van de mate van participatie door de schoolleider wordt rechtstreeks aangesloten bij het onderzoek naar isolement van leerkrachten (Bakkenes e.a., o.c.), omdat in een communicatienetwerk alle vaste leden van het schoolteam zijn opgenomen, ongeacht hun functie. Met het oog op de vergelijkbaarheid van het onderzoek naar isolement bij leerkrachten en bij schoolleiders wordt de vragenlijst naar taakperceptie bij schoolleiders zoveel mogelijk analoog opgezet aan de vragenlijst naar taakperceptie bij leerkrachten. De schoolleiders die aan het onderzoek deelnemen vullen twee vragenlijsten in: een vragenlijst voor de taakperceptie en een vragenlijst over commu-

nicatie met teamleden.

Een absolute voorwaarde voor het samenstellen van goed interpreteerbare communicatienetwerken is de deelname van volledige schoolteams aan het onderzoek. Dit is een complicerende factor bij het werven van scholen, omdat de weigering van een kleine minderheid binnen een schoolteam om aan het onderzoek mee te werken deelname van de hele school aan het onderzoek verhindert. Naast elkaar zijn diverse strategieën gevolgd om een voldoende omvangrijke proefgroep van scholen (ongeveer 30) samen te stellen, zoals samenwerking met schoolbegeleidingsdiensten die in het onderzoek geïnteresseerd zijn, en bekendheid van de onderzoeker met scholen in een bepaalde regio. Aan de eis van representativiteit van de steekproef ten opzichte van de landelijke populatie wordt daarmee niet voldaan. Wel is gestreefd naar een groep scholen die evenwichtig is samengesteld wat betreft interne collegiale verhoudingen, denominatie, grootte en urbanisatie. Gegeven het eigen karakter van traditionele vernieuwingsscholen met het oog op interne afstemming van onderwijskundige zaken worden deze buiten de proefgroep gehouden.

3.2 Instrumenten

Taakperceptie

Het instrument om taakperceptie bij schoolleiders te meten is op vergelijkbare wijze samengesteld als het instrument voor taakperceptie bij leerkrachten (Koning & De Water, 1994). Overwegingen daarbij zijn dat het leerkrachten-instrument goed interpreteerbare gegevens heeft opgeleverd, en dat de vergelijkbaarheid van de resultaten wordt bevorderd. Het leerkrachten-instrument bestaat uit drie vragen over de taakkenmerken autonomie, variatie en feedback:

- hoeveel zeggenschap heeft U over de uitvoering van deze taak?
- hoeveel variatie kent de taak?
- in hoeverre biedt de uitvoering van deze taak op zichzelf informatie over de kwaliteit van uw functioneren (als schoolleider)?

De schoolleiders beantwoorden dezelfde drie vragen voor 24 taken; de respondent geeft op 7-puntsschalen aan in welke mate zij de drie hiervoor beschreven taakkenmerken van toe-

passing vinden op de 24 taken. Bij het samenstellen van een takenlijst voor schoolleiders is uitgegaan van het onderscheid in drie taakdomeinen (par 2.3). Tevens is rekening gehouden met het gegeven dat het merendeel van de directeuren van basisscholen een substantiële lesgevende taak heeft. Met betrekking tot deze taken wordt geen verschil verwacht in taakperceptie tussen geïsoleerde en niet-geïsoleerde schoolleiders, in tegenstelling tot onderwijskundige leiderschapstaken. Om die redenen is een aantal instructietaken en voortgangsbewakingstaken uit het professionele domein in de takenlijst voor schoolleiders opgenomen. Daarnaast is gebruik gemaakt van gangbare a priori indelingen van schoolleiderstaken in categorieën, die te vinden zijn in functie- en taakanalyses (Van Greevenbroek, 1981; Imants, 1986; Gemeente Rotterdam, 1990; De Wit, 1991). Deze a priori indelingen van taken zijn gebruikt voor het samenstellen van een bondige en tevens representatieve takenlijst. Uiteindelijk is een lijst samengesteld van 24 taken die in random volgorde aan de respondenten worden aangeboden. De 24 taken zijn verdeeld over zeven categorieën¹:

- verzorgen van opvoeding en onderwijs (5 taken; professioneel domein)
- planning en beleid (5 taken; overlap gebied)
- beheer (5 taken; beheers- en bestuursdomein)
- interne communicatie (2 taken; overlap gebied)
- externe contacten (4 taken; beheers- en bestuursdomein)
- evaluatie (2 taken; overlap gebied)
- professionalisering (1 taak; overlap gebied).

Participatie in het communicatienetwerk

De mate van taakgerichte communicatie door schoolleiders wordt vastgesteld op basis van een communicatienetwerk van de school waarin de schoolleider werkzaam is. Voor de constructie van een dergelijk netwerk vullen alle leden van een schoolteam een vragenlijst in waarin gevraagd wordt naar de frequentie van taakgerichte communicatie met de overige teamleden (geschat op een 7-puntsschaal tussen 'nooit' en 'dagelijks'), en naar het belang dat aan deze taakgerichte communicatie wordt toegekend (geschat op een 4-puntsschaal tus-

sen 'niet van belang' en 'van groot belang'). Participatie wordt vervolgens geoperationaliseerd als het verschil tussen het produkt van de rijssommen van belang en frequentie en het produkt van het aantal deelnemers in het netwerk en de ondergrens van communicatie binnen het netwerk (rijssommen van belang $\times$ frequentie - $n \times$ ondergrens)².

3.3 Data-analyse

Op de taakperceptiegegevens werd allereerst een principale componentenanalyse (PCA) uitgevoerd. In deze analyse werden de percepties van de 24 taken per respondent opgevat als onafhankelijke observaties, zodat een gegevensbestand ontstond van $N \times 24$ observaties. De bedoeling van deze analyse was niet om tot datareductie te komen, maar om een ruimtelijk model te ontwikkelen van de observaties in termen van de drie taakperceptievariabelen waarin de totale hoeveelheid variantie behouden bleef. Dit model diende vervolgens als standaard waarbinnen de relaties tussen de mate van participatie in het communicatienetwerk en de taakperceptie bij verschillende categorieën van taken zichtbaar gemaakt konden worden. Immers in de PCA krijgen alle observaties een positie in de principale componentenruimte, en wel zo dat observaties dicht bij elkaar geplaatst worden naarmate hun profielen op de drie taakkenmerken meer met elkaar overeenkomen. Dit opent dan de mogelijkheid om de gemiddelde positie per taak te berekenen en te onderzoeken in hoeverre de a priori indeling van de taken is terug te vinden in de posities die taken innemen in de principale componentenruimte. Daarmee wordt vastgesteld in hoeverre de perceptie van de taken en de veronderstelde motivatie voor deze taken die deel uitmaken van eenzelfde categorie vergelijkbaar is.

Om de relatie tussen netwerkparticipatie en taakperceptie binnen verschillende categorieën van taken te onderzoeken werd per categorie een canonische correlatie berekend tussen netwerkparticipatie en de component-scores uit de PCA-oplossing. In feite werd dus voor elke categorie van taken een richting in de PCA-ruimte gefit die het verband tussen participatie en taakperceptie beschrijft. Uit de stand van deze richting ten opzichte van de vectoren voor de taakperceptievariabelen kan dan wor-

den afgeleid hoe de netwerkparticipatie zich verhoudt tot de drie aspecten van taakperceptie. Vanwege de duidelijke verschillen in grootte tussen de scholen werd op dezelfde wijze ook de relatie tussen de grootte van het netwerk (= de omvang van de school als organisatie) en taakperceptie onderzocht. Voor elke categorie van taken apart werd dus nagegaan in hoeverre het aantal leerkrachten dat deel uitmaakt van het netwerk samenhangt met de taakperceptie. De PCA werd uitgevoerd met PRINCALS uit het SPSS-pakket. Voor de canonische correlaties werd de GLM-procedure uit het SAS-pakket gebruikt.

Hoewel beschrijving van sociogrammen op basis van de communicatienetwerken voor de beantwoording van de vraagstelling strikt genomen niet noodzakelijk is, worden in paragraaf 4 enkele sociogrammen gepresenteerd, omdat deze het verschil in positie van schoolleiders tussen scholen goed zichtbaar maken. Een sociogram van een school wordt ontwikkeld op basis van een dichotomisering van de communicatiegegevens. Uitgaande van de gedachte dat er sprake is van een contact, wanneer twee teamleden beiden aangeven tenminste wekelijks (schaalwaarde 5) met elkaar over het werk te praten en dit beiden belangrijk te vinden (schaalwaarde 3), wordt voor beide deelnemers een score van 15 gehanteerd als grenswaarde voor het produkt tussen frequentie en belang (1 = wel communicatie; 0 = geen communicatie). Een netwerk geeft de onderlinge relaties tussen teamleden van een school weer op basis van alle aldus vastgestelde 1-scores.

4 Resultaten

4.1 Beschrijving van de proefgroep

Uiteindelijk hebben 23 basisscholen (12 openbaar, 6 P.C., 4 R.K. en 1 neutraal bijzonder) deelgenomen aan het onderzoek. Deze scholen zijn gelegen in het zuidwesten en het midden van Nederland. Wat betreft de interne colle-

giale verhoudingen beschikken we niet over 'harde' gegevens. In overleg met de betrokken schoolbegeleidingsdiensten is gestreefd naar een evenwichtige verdeling van scholen in de proefgroep. Wat betreft grootte, denominatie en urbanisatie is de verdeling als volgt. De grootte van de scholen varieert tussen een school met 5 teamleden en een school met 17 teamleden. De grootte is regelmatig gespreid, waarbij de teamgroottes 6, 9, 10 en 14 ieder driemaal voorkomen. Op alle scholen op één na zijn een of meerdere part timers werkzaam, variërend van 4,5 tot 36 uur. Vakleerkrachten, OETC-leerkrachten en remedial teachers zijn in dit overzicht niet opgenomen. Twaalf scholen bevinden zich in een stad, de overige elf scholen staan in dorpen met een uitgesproken plattelandskarakter of in dorpen die verstedelijken. Twee scholen uit de proefgroep beschikken over tweehoofdig directeurschap. Van de 23 scholen hebben 25 directeurs de vragenlijsten voldoende ingevuld. Over de gegevens van deze 25 schoolleiders wordt in het nu volgende gerapporteerd.

4.2 Taakperceptie

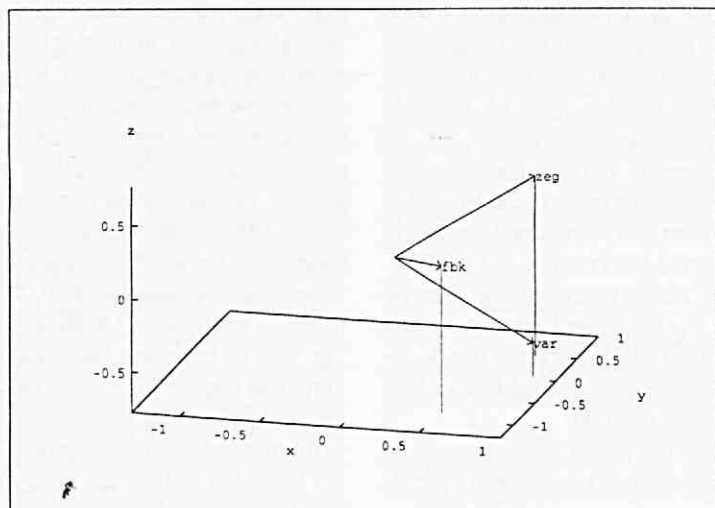
Door alle 24 taken als onafhankelijke observaties op te vatten ontstaat een bestand van 600 (25×24) observaties. Omdat de analyse een ruimtelijk model moet opleveren waarin de oorspronkelijke variantie behouden blijft is voor de drie perceptievariabelen een driedimensionale oplossing gekozen. De eigenwaarden van deze oplossing bedragen .50, .32 en .19. De componentladingen zijn weergegeven in Tabel 1 en Figuur 1.

De eerste dimensie in de Figuren 1 en 2 (x) beschrijft wat de drie variabelen gemeenschappelijk hebben. Volgens Hackman en Lawler (o.c.) is de motivatie sterker naarmate taken hoger gewaardeerd worden in termen van de drie taakperceptievariabelen. Elk van deze variabelen levert een bijdrage aan de betekenisvolheid van het werk. Door de substantiële positieve lading van elk van de drie variabelen op de eerste

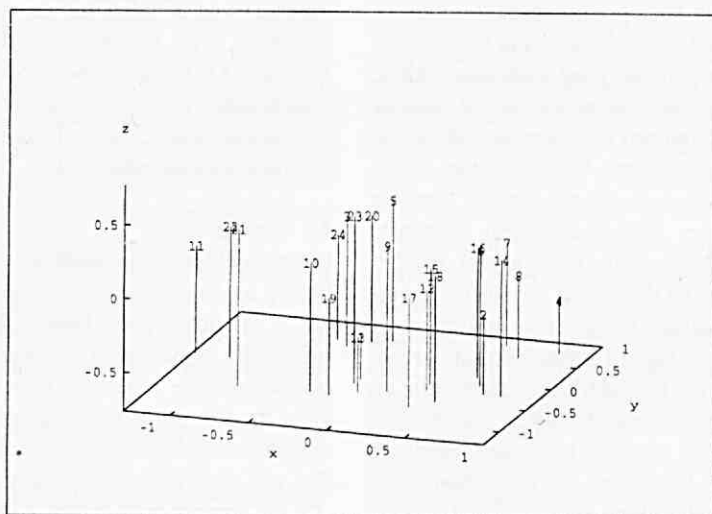
Tabel 1

Componentladingen van de drie taakperceptievariabelen in een drie-dimensionale PCA-oplossing

	dimensie 1	dimensie 2	dimensie 3
variatie	.839	.071	-.545
autonomie	.737	.499	.460
feedback	.513	-.831	.230



Figuur 1. Componentladingen van de taakperceptievariabelen in een driedimensionale ruimte



Figuur 2. Centroïden van de 24 taken in de driedimensionale componentenruimte

dimensie lijkt de interpretatie van de drie variabelen als motivationele variabelen ondersteund te worden. De eerste dimensie kan worden opgevat als taakmotivatie. Op de tweede dimensie (y) wordt in Figuur 1 feedback onderscheiden van autonomie en variatie. De derde dimensie (z) van Figuur 1 toont een onderscheid tussen variatie enerzijds en autonomie en feedback – zij het in mindere mate – anderzijds. Dit betekent dat de drie taakperceptie-variabelen duidelijk van elkaar onderscheiden kunnen worden, maar tevens een gemeenschappelijke component hebben.

De drie vectoren in Figuur 1 zijn ongeveer even lang, zoals uit de componentladingen

voor de drie dimensies (Tabel 1) kan worden afgeleid (een perspectief-vertekening werkt de kortere indruk van de feedback-vector in Figuur 1 in de hand), en de drie variabelen zijn gelijkmatig verspreid in de ruimte. De positie van een observatie op de eerste dimensie geeft aan in hoeverre de betreffende taak door een respondent als motiverend wordt waargenomen. Afwijkingen van het nulpunt op de tweede en derde dimensie geven weer in hoeverre de variabelen autonomie, variatie en feedback in hun bijdrage aan taakmotivatie verschillen. In Figuur 2 zijn, gemiddeld over alle respondenten, de centroïden van de 24 taken in de drie dimensionale ruimte weergege-

Tabel 2*Categorieën van schoolleiderstaken in relatie tot taakdomeinen en taakperceptie*

categorie	domein	perceptie	taaknummer en taak
instructie en begeleiding	1	hoog	4. instructie geven aan individuele of kleine groepjes leerlingen 7. instructie geven aan de hele groep 8. met een leerling zijn/haar ontwikkeling, prestaties en vorderingen bespreken
creëren van een open schoolcultuur	2	hoog	2. representatie van de school naar de omgeving 6. bijhouden van ontwikkelingen op het gebied van onderwijs en schoolmanagement 14. het stimuleren van contacten tussen teamleden over onderwijsinhoudelijke en didactische zaken 16. het actief bij de school betrekken van ouders
onderwijskundige afstemming en coördinatie	2	meer dan gemiddeld	9. met collega's tot een verdeling komen van op het onderwijs gerichte taken 12. het voeren van personeelsbeleid: loopbaanbegeleiding, nascholing, voortgezette scholing 15. met collega's bespreken welke leermiddelen, materialen e.d. worden aangeschaft 17. uitvoeren van evaluaties m.b.t. schoolkwaliteit 18. met teamleden de uitkomsten en consequenties van evaluaties bespreken
voorwaarden voor het onderwijskundig functioneren van de school	2/3	minder dan gemiddeld	1. ontwikkelen schoolwerkplan 10. opzetten en bijhouden van een leerlingvolgsysteem 13. werven en selecteren van personeel 19. het leiding geven aan vergaderingen en andere vormen van overleg- en besluitvormingsprocessen 23. het onderhouden van contacten met externe instanties (bestuur, inspectie, etc.)
administratieve taken	3	laag	11. uitvoeren van administratieve verrichtingen m.b.t. het leerlingenbestand 21. het beheren van de financiële middelen (F.B.S.) en de materiële voorzieningen (gebouw, inventaris, goederen) 22. uitvoeren van administratieve verrichtingen m.b.t. het personeelsbestand
overige taken	1 3	minder dan gemiddeld minder dan gemiddeld	5. beoordelen en corrigeren van het werk van leerlingen 20. organiseren en begeleiden van sport- en speldagen/ festiviteiten 3. het schrijven van het jaarverslag 24. het verzamelen en verspreiden van informatie van buiten de school

ven (voor de bij de taaknummers behorende taken wordt verwezen naar Tabel 2). Uitgaande van de houdbaarheid van de hiervoor beschreven interpretatie van de drie-dimensionale ruimte, blijkt uit Figuur 2 dat de onderzochte schoolleiders sterk gemotiveerd zijn voor taken op het gebied van onderwijs en opvoeding, en zwak gemotiveerd voor strikt administratieve taken (het betreft hier de spreiding van de taken langs de x-as).

Een inhoudelijke beoordeling van de taken in combinatie met de resultaten van beide ana-

lyses leidt tot de indeling in categorieën volgens Tabel 2. De PRINCALS-analyse en een aanvullende cluster-analyse, waarover hier niet nader wordt gerapporteerd, geven aanleiding om vijf categorieën van taken te onderscheiden, naast vier taken die in een restcategorie zijn ondergebracht. Vier van de vijf categorieën kunnen worden gekarakteriseerd aan de hand van de afstand van taken tot het werk met leerlingen, waarbij instructie- en begeleidingstaken direct op het werk met leerlingen betrekking hebben, en administratieve ta-

ken de grootste afstand tot het werken met leerlingen kennen. Enigszins afwijkend van deze lijn zijn taken in de tweede categorie die relatief sterk met elkaar clusteren, en die een hoge taakmotivatie kennen. Gemeenschappelijk aan deze taken is dat ze gezien kunnen worden als behorende tot een categorie gericht op het creëren van een open cultuur in de school voor interne en externe uitwisseling van informatie.

In het algemeen geldt dat schoolleiders de hoogste perceptie-scores te zien geven voor direct op het werken met leerlingen gerichte taken, ofwel taken uit het professionele taakdomein (domein 1). De perceptiescores nemen af als de afstand tot het werken met leerlingen groter wordt. De taken uit het beheers- en bestuursdomein (domein 3) vertonen de laagste perceptie-scores, en taken uit het gebied waar beide domeinen elkaar overlappen (domein 2) scoren rond het gemiddelde. De taken die betrekking hebben op het creëren van een open cultuur (domein 2) vertonen eveneens hoge perceptie-scores.

4.3 Netwerkp participatie en communicatienetwerken

Wanneer de mate van participatie van schoolleiders binnen de netwerken van hun scholen wordt berekend op de in 3.3 besproken wijze levert dat de volgende resultaten op. De netwerkp participatie varieert tussen 867 (laagste participatie) en 3402 (hoogste participatie), met als gemiddelde participatie 1929 en standaarddeviatie 637. De resultaten geven aan dat van een evenwichtige spreiding van mate van participatie gesproken kan worden. Om de participatiescores een meer concrete betekenis te geven beschrijven we hier drie communicatienetwerken, mede ter illustratie van het volgende:

- het verschil tussen relatief hechte en losse netwerken,
- het verschil in mate van communicatie van leden binnen één netwerk,
- het verschil tussen hechte en losse koppeling van de directeur in het netwerk.

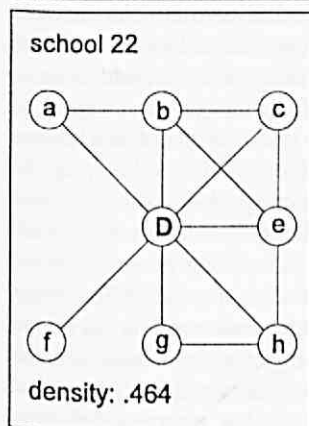
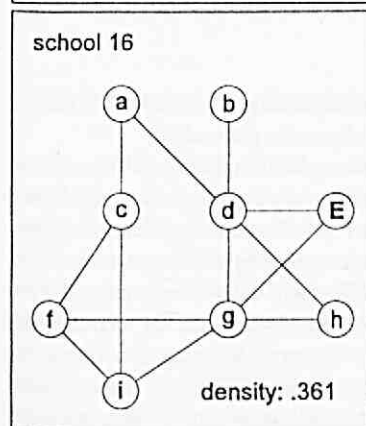
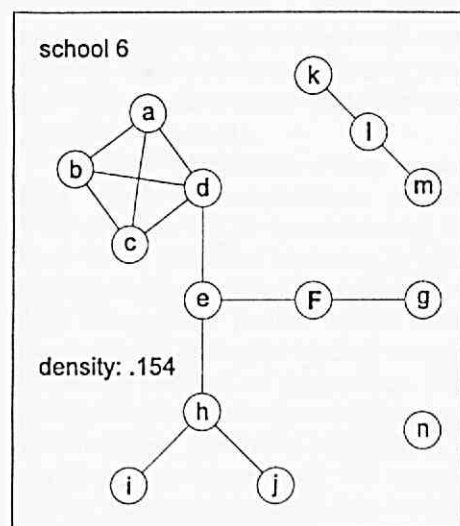
Figuur 3 toont een los netwerk (density = .154) en twee hechte netwerken (density = .361 en .464). In het losse netwerk van school 6 treffen we een volledig geïsoleerde leerkracht aan (N) en twee leerkrachten (K en M) die slechts door

een derde leerkracht (L) communicatie vertonen. Daarnaast maakt een subgroep van vier leerkrachten deel uit van het netwerk (A t/m D). Deze groep is door slechts een enkele leerkracht (D) verbonden met de rest van het netwerk langs een leerkracht in de rol van verbindingspersoon (E). De schoolleider (F) heeft met twee leerkrachten een verbinding (E en G), en verkeert daarom in een naar isolement neigende positie (participatie score 1319). In het hechte netwerk van school 22 zijn alle teamleden op sterke of zwakke wijze aan collega's gekoppeld. Een groep van vier teamleden kan onderscheiden worden (B t/m E), waarvan de directeur (D) deel uitmaakt. De directeur is de enige deelnemer in het netwerk die met alle collega's communiceert (participatie score 3033), en functioneert daarom als verbindingspersoon tussen de groep en teamleden die niet van de groep deel uitmaken (A, F, G en H). Een leerkracht (F) is slechts door de directeur met het netwerk verbonden, en neemt daarom een naar isolement neigende positie in.

Uit de bespreking van de scholen 6 en 22 zou de suggestie gewekt kunnen worden dat de dichtheid van het netwerk samengaat met de mate van communicatie van de directeur. Om aan te geven dat dit niet het geval hoeft te zijn is het netwerk van school 16 opgenomen. Deze school kent een relatief hecht netwerk, waarin twee clusters van vier teamleden op één verbinding na groepen vormen (C, F, G en I; D, E, G en H). Twee leerkrachten (D en G) nemen verbindingsposities in, waarbij G de twee bijna-groepen verbindt en D twee relatief geïsoleerde leerkrachten (A en B) aan het netwerk bindt. Directeur E maakt alleen van een bijna-groep deel uit, en heeft daarbinnen slechts twee verbindingsposities. Deze directeur neemt een relatief geïsoleerde positie in binnen een hecht netwerk (participatie score 1400).

4.4 Netwerkp participatie en taakperceptie

Met deze analyse, die gericht is op de beantwoording van de onderzoeksvraag, gaat het erom aanwijzingen te vinden voor de relatie tussen netwerkp participatie en taakperceptie per categorie van taken. De aantallen observaties waarover we in het kader van dit onderzoek beschikken leidt ertoe dat we moeten spreken van een eerste explorerende analyse die richtingen kan aangeven voor verder onderzoek. Voor



Tabel 3

Resultaten canonische correlatie-analyses per categorie van taken

	r_{cn}	w_{c1}	w_{c2}	w_{c3}
instr	.142	.835	.597	.149
open	.312*	.936	-.813	.588
coor	.257	1.147	-.712	.223
voor	.237*	.346	-1.016	.232
adm	.195	-.422	-.257	-.791

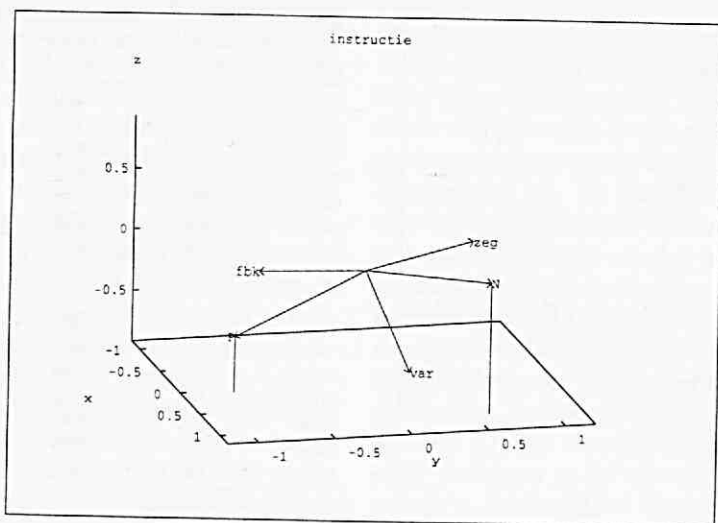
	r_{cp}	w_{c1}	w_{c2}	w_{c3}
instr	.177	.054	-.861	-.468
open	.104	.958	.430	.742
coor	.250	.947	-.755	.508
voor	.250*	.859	-.473	.485
adm	.181	.427	.505	-.797

1 en 3. Voor elke categorie van taken is een canonische correlatie berekend tussen netwerkparticipatie en de principale componenten in het ruimtelijke model van de taakpercepties. Dat betekent dat we in dit principale componentenmodel een richting fitten die maximaal correleert met netwerkparticipatie. Uit de stand van deze richting kan dan afgeleid worden hoe netwerkparticipatie zich verhoudt tot de drie oorspronkelijke aspecten van taakperceptie. Deze analyse is per categorie van taken uitgevoerd door de analyse te beperken tot die observaties die betrekking hebben op de taken die deel uitmaken van een categorie. Omdat we bij het zoeken naar een adequate operationalisatie van netwerkparticipatie op de invloed van de grootte van het netwerk waren gestuit, zijn we op de dezelfde wijze ook de relatie tussen taakperceptie en netwerkgrootte nagegaan. Het resultaat van deze analyses is weergegeven in Tabel 3. In deze Tabel 3 zijn per taakcategorie weergegeven de canonische correlatie tussen netwerkgrootte respectievelijk netwerkparticipatie en de gewichten voor de drie componenten die de canonische richting in de principale componentenruimte beschrijven. Het bovenste gedeelte van de tabel heeft betrekking op de netwerkgrootte, het onderste op de netwerkparticipatie.

Het geheel aan uitkomsten overziende zijn er interpreteerbare verschillen in sterkte van de samenhang met taakperceptie. Bij enkele categorieën bereikt de canonische correlatie een significant niveau, hetgeen is aangegeven met een asterisk. In die gevallen is meer dan 5% van

Figuur 3. De communicatienetwerken van drie scholen

taakcategorieën uit domein 2 verwachten we wel verschillen in taakperceptie tussen geïsoleerde en niet-geïsoleerde leerkrachten, in tegenstelling tot taakcategorieën uit de domeinen



Figuur 4. Canonische richting in de principale componentenruimte voor netwerk grootte (N) en netwerk participatie (P) bij 'instructie' taken

de variantie van de voorspellende variabele gerelateerd aan taakperceptie. De aard van de samenhang kan ook grafisch worden weergegeven, zoals bijvoorbeeld is gebeurd in Figuur 4³. In vergelijking met Figuren 1 en 2 is deze figuur zodanig gedraaid dat we nu tegen de tweede dimensie (y) aankijken in plaats van tegen de eerste (x), omdat deze draaiing een beter zicht verschaft. Bij de taken die betrekking hebben op het verzorgen van opvoeding en onderwijs (instructie) lijkt er geen sprake van een substantiële relatie, noch met de netwerk grootte, noch met de netwerk participatie. Voor zover we echter, de correlatie met de netwerk participatie nog enige waarde zouden mogen toekennen, zou deze gegeven de stand van de canonische richting (vector P in Figuur 4) erop duiden dat schoolleiders die gekenmerkt worden door een grotere participatie meer feedback en minder zeggenschap toeschrijven aan instructietaken, terwijl schoolleiders met kleinere participatie juist minder feedback en meer zeggenschap aan deze taken toekennen. Per saldo is er dan geen verschil in perceptie als zodanig.

Bij taken die gericht zijn op het bevorderen van een open cultuur is er alleen een samenhang met de netwerk grootte die bovendien een significant niveau bereikt. De canonische richting laat zien dat deze taken in grotere scholen in het algemeen positiever waargenomen worden, vooral omdat ze feedback leveren. Met

deze samenhang is ongeveer 9% van de variantie in netwerk grootte gebonden.

Bij de taken die betrekking hebben op onderwijskundige afstemming en coördinatie (coördinatie) wordt geen significante samenhang gevonden, maar het niveau van de samenhang is van dien aard dat we vermoeden dat zowel participatie als grootte in een minder exploratieve onderzoeksofzet van betekenis zullen blijken te zijn. Kijken we onder die aanname naar de aard van de samenhang, dan blijkt dat deze taken positiever waargenomen worden in grotere netwerken en door schoolleiders met een grotere graad van participatie. Deze samenhang komt in beide gevallen vooral tot stand omdat deze taken meer feedback leveren.

Bij taken die betrekking hebben op het scheppen van voorwaarden voor het onderwijskundig functioneren van de school wordt een significante samenhang gevonden zowel met netwerk grootte als met netwerk participatie. De significantie vloeit voort uit het feit dat deze categorie een groter aantal taken kent dan de andere. Schoolleiders met een hoge graad van participatie vinden "voorwaardenscheppende taken attractiever dan schoolleiders met een lagere participatiegraad en ze leggen daarbij een licht accent op de feedback die ze op basis van deze taken krijgen. De aard van de samenhang ligt bij de netwerk grootte iets anders. In grotere netwerken worden voorwaardenscheppende taken slechts in geringe mate

positiever waargenomen. In grotere netwerken nemen schoolleiders deze taken vooral waar als taken die feedback leveren. In kleinere netwerken wordt de waarneming bepaald door de zeggenschap die ze bieden.

Wat betreft administratieve taken worden er geen significante samenhangen gevonden. Voor zover we aan de correlaties toch betekenis willen toekennen impliceert de canonische richting dat op grotere scholen administratieve taken minder positief worden waargenomen, vooral omdat het bij deze taken ontbreekt aan feedback en zeggenschap. Bij de participatiegraad zou het verband inhouden dat administratieve taken door schoolleiders met een hogere participatiegraad wat positiever worden waargenomen, omdat ze in het takenpakket afwisseling bieden.

5 Conclusies en discussie

Vanwege beperkingen in de samenstelling van de proefgroep wat betreft omvang en representativiteit moeten de conclusies van dit onderzoek met voorzichtigheid worden geformuleerd. De gevonden verschillen tussen geïsoleerde en niet-geïsoleerde schoolleiders kunnen worden gezien als tendenties, en zijn in een grotere proefgroep wellicht duidelijker aantoonbaar. De gevonden tendensen wijzen hoe dan ook in de verwachte richting. Dat betekent dat geïsoleerde en niet-geïsoleerde schoolleiders geen duidelijk verschil vertonen in taakperceptie van lesgevende taken en algemene beheersmatige taken. Voor zover verschillen gevonden zijn tussen schoolleiders met een hoge en een lage mate van participatie doen deze zich voor bij taken die betrekking hebben op onderwijskundig leiderschap, met name coördinatie van onderwijs en het scheppen van voorwaarden voor het onderwijs: schoolleiders met een hoge mate van participatie nemen taken betreffende coördinatie van en voorwaarden voor het onderwijs positiever waar dan schoolleiders met een lage mate van participatie.

Schoolleiders van de in dit onderzoek betrokken basisscholen vertonen verschil in de mate van participatie in het communicatienetwerk van hun scholen. Daarmee is een aanwijzing gevonden voor variatie in mate van isole-

ment van directeurs van basisscholen ten opzichte van hun teamleden. Een kanttekening bij deze conclusie is dat in dit onderzoek de grootte van het communicatienetwerk wellicht een rol speelt bij de mate van communicatie door de directeur met teamleden. Bij het vaststellen van de mate van participatie is een correctie toegepast voor de grootte van het netwerk. Opmerkelijk is dat de grootte van het netwerk als zodanig samenhang vertoont met de waarneming van taken door schoolleiders. Door Stoel (o.c.) is geen samengang gevonden tussen de grootte van de school en de taken die door schoolleiders in basisscholen worden uitgevoerd. Toekomstig onderzoek moet over de exacte aard van deze relaties meer duidelijkheid verschaffen.

De schoolleiders in dit onderzoek is gevraagd om een breed scala aan schoolleiderstaken te waarderen in termen van zeggenschap, variatie en feedback. Op basis van deze waarderingsen kan een indeling worden gemaakt in categorieën van taken, die op een dimensie geplaatst kunnen worden. Directeuren van de onderzochte basisscholen nemen taken op het terrein van onderwijs en opvoeding waar als de meest aantrekkelijke taken en strikt administratieve taken als de minst aantrekkelijke taken. Wanneer de taken verder van het werk met leerlingen verwijderd raken worden deze als minder aantrekkelijk waargenomen; van instructietaken via taken gericht op de coördinatie van het onderwijs en het scheppen van voorwaarden voor het onderwijs naar strikt administratieve taken. Enigszins afwijkend van deze dimensie zijn de als aantrekkelijk waargenomen taken betreffende het creëren van een open cultuur in de school. Kenmerkend voor deze categorie is de aandacht voor het open stellen van communicatiekanalen in en rond de school voor informatie over het onderwijs. Wanneer we een vergelijking maken met isolement en taakperceptie bij leerkrachten (Bakkenes e.a., o.c.), dan valt in het oog dat isolement en taakperceptie bij leerkrachten en schoolleiders in elkaars verlengde liggen. Redenen voor deze conclusie zijn (1) dat beide groepen de dimensie 'afstand ten opzichte van het werken met leerlingen' hanteren om taken waar te nemen, (2) dat voor beide groepen geldt dat motivatie voor taken gericht op het werken met leerlingen het hoogst is, en (3) dat verschil

tussen geïsoleerde en niet-geïsoleerde leerkrachten én schoolleiders zich met name voor- doet bij de taken die vallen binnen het gebied waar het professionele domein en het bureau- cratische domein van de school elkaar overlap- pen. Geredeneerd vanuit de drie eerder in dit artikel besproken taakdomeinen vormen isole- ment van leerkrachten en van schoolleiders als het ware elkaars spiegelbeeld, althans wanneer de lesgevende taak van de schoolleider buiten beschouwing wordt gelaten. Peters en Slegers (1994) plaatsen een vraagteken bij het spiege- len van de professionele oriëntatie van leer- krachten en schoolleiders op basis van de drie taakdomeinen. Zij stellen dat algemene mana- gementtaken en onderwijskundige leider- schapstaken in de functie van schoolleider gecombineerd moeten worden. We onder- schrijven het belang van dit uitgangspunt voor een doelmatige invulling van de schoolleiders- functie. Dit laat onverlet dat de resultaten van het onderzoek naar isolement en de feitelijke taakperceptie van schoolleiders in basisscho- len wijzen in de richting van isolement van leerkrachten en schoolleiders als elkaars spie- gelbeeld.

Door Stoel (o.c.) is gevonden dat veel schoolleiders in basisscholen onderwijskun- dige leiderschapstaken als de kern van hun be- roep zien, waar zij door de druk van algemene managementtaken slechts in beperkte mate aan toekomen. In het hier beschreven onderzoek hebben de deelnemende schoolleiders een po- sitiever waarneming van onderwijskundige lei- derschapstaken dan van algemene manage- menttaken. Tot zover worden de uitkomsten van het onderzoek van Stoel bevestigd. Een be- langrijke aanvulling is dat schoolleiders onder- ling verschil vertonen in de wijze waarop zij deze taken waarnemen, en dat de min of meer geïsoleerde positie van de schoolleider binnen het schoolteam daarbij mogelijk een belang- rijke rol speelt. Bovendien lijken alle schoollei- ders in het onderzoek het sterkst gemotiveerd te zijn voor taken gericht op het directe werk met leerlingen. Echter, schoolleiders die het accent leggen op de uitvoering van direct op instructie en begeleiding van leerlingen gerichte taken worden doorgaans niet met onderwijskundig leiderschap in verband gebracht (zie bijv.: Ful- lan, 1991).

Noten

1. De 24 taken zijn opgenomen in Tabel 2; in deze tabel is de a priori indeling vervangen door de in- deling op basis van de respondentscores.
2. In eerste opzet was beoogd om de zogenaamde genormaliseerde netwerkparticipatie te gebrui- ken. In deze maat wordt voor de grootte van het netwerk gecorrigeerd door de participatie uit te drukken als proportie van de maximaal moge- lijke score. De resulterende waarde bleek echter sterk negatief te correleren met de grootte van het netwerk. Verschillende oorzaken kunnen ten grondslag hebben gelegen aan dit sterk nega- tieve verband met de grootte van het netwerk, waarop nu verder niet wordt ingegaan. Een meer adequate correctie hebben we daarom gezocht in het aftrekken van een van de grootte van het netwerk afhankelijke ondergrens van communi- catie. Deze keuze is tot stand gekomen op basis van de redenering dat iedereen wel met iedereen een minimum aan communicatie onderhoudt, zodat de totale hoeveelheid communicatie af- hankelijk is van de grootte van het netwerk. Leg- gen we deze 'drempel' aan, dan blijft de belang- rijke, en van de grootte van het netwerk onafhankelijke communicatie over. Concreet is de grens gelegd bij de waarde 9 als uitkomst van het produkt tussen frequentie en belang. Met K participanten in het netwerk is $K \times 9$ dan de per netwerk af te trekken hoeveelheid communica- tie. Deze maat bleek slechts licht positief (.29) met de grootte van het netwerk te correleren.
3. Voor elke taakcategorie is een vergelijkbare fi- guur samengesteld. Met het oog op de beperkte ruimte wordt hier alleen Figuur 4 afgebeeld.

Literatuur

- Alblas, G., & Vliert, E. van de (1990). *Doelmatig orga- niseren*. Groningen: Wolters-Noordhoff.
- Ashforth, B., & Lee, R. (1990). Defensive behavior in organizations: A preliminary model. *Human Re- lations*, 43, 621-648.
- Ashton, P., & Webb, R. (1986). *Making a difference: Teachers' sense of efficacy and student achieve- ment*. New York: Longman.
- Bakkenes, I., Brabander, K. de, & Imants, J. (1994). Het taakspecifieke karakter van professioneel isolement in basisscholen. *Pedagogisch Tijd- schrift*, 19, 381-409.

- Bakkenes, I., Imants, J., & Brabander, C. de (1994). Professioneel isolement van leerkrachten in het basisonderwijs. In F. Korthagen, T. Wubbels, & T. Bergen (Red.), *VELON-Jaarboek: Recent onderzoek over onderwijs en opleiding* (pp. 15-34). Utrecht/Groningen: WCC/VELON.
- Barth, R. (1987). *School: A community of leaders*. Paper presented at the Annual spring Conference of the Georgia State University Principals (ERIC Doc. 281 277).
- Commissie Evaluatie Basisonderwijs (1994). *Zicht op kwaliteit: evaluatie van het basisonderwijs*. De Meern: Inspectie van het Onderwijs.
- Daft, R., & MacIntosh, N. (1981). A tentative exploration into the amount and equivocality of information processing in organizational work units. *Administrative Science Quarterly*, 26, 207-224.
- Dallinger, J. (1985, november). *The influence of task characteristics upon task-communicating network roles*. Paper presented at the 71st annual meeting of the Speech Communication Association, Denver.
- Farace, R., Monge, P., & Russell, H. (1977). *Communicating and organizing*. Reading, MA: Addison-Wesley.
- Firestone, W., & Wilson, B. (1985). Using bureaucratic and cultural linkages to improve instruction: the principal's contribution. *Educational Administration Quarterly*, 21(2), 7-30.
- Flinders, D. (1988). Teacher isolation and the new reform. *Journal of Curriculum and Supervision*, 4(1), 17-29.
- Fullan, M. with Stiegelbauer, S. (1991). *The new meaning of educational change*. London: Cassell.
- Gemeente Rotterdam (1990). *Instructie, functie- en taakomschrijving onderwijspersoneel openbaar basis- en (voortgezet) speciaal onderwijs*.
- Greevenbroek, A. van (1981). *Onderwijsassistent en taakdifferentiatie*. 's Gravenhage: Staatsuitgeverij.
- Hackman, J., & Lawler, E. (1971). Employee reactions to job characteristics. *Journal of Applied Psychology*, 55, 259-286.
- Hanson, E. (1979). *Educational administration and organizational behavior*. Boston: Allyn and Bacon.
- Imants, J. (1986). *Taakdifferentiatie en besluitvorming in basisscholen*. Den Haag: SVO.
- Imants, J. (1989). Taakdifferentiatie en doelmatigheidsbeleving in schoolorganisaties. In B. Creemers e.a. (Eds.), *Handboek schoolorganisatie en onderwijsmanagement* (pp. 2120-1-2120-17). Alphen a/d Rijn: Samsom.
- Imants, J. (1994). Taakdifferentiatie en motivatie van leraren. In B. Creemers e.a. (Eds.), *Handboek schoolorganisatie en onderwijsmanagement* (pp. A 1025-1-A 1025-23). Alphen a/d Rijn: Samsom.
- Imants, J., & Bakkenes, I. (1993). Professional isolation of teachers, school culture and teachers' sense of efficacy. In F. Kieviet & R. Vandenbergh (Eds.), *School culture, school improvement and teacher development* (pp. 151-169). Leiden: DSWO Press.
- I.T.S. (1975). *Functies en taken van onderwijsgeven-den*. Nijmegen: I.T.S.
- Koning, I., & Water, P. de (1994). *Professioneel isolement van schoolleiders in het basisonderwijs*. Leiden: Rijksuniversiteit / Vakgroep Onderwijsstudies (doctoraalscriptie).
- Lindle, J., & De Marco, L. (1988). *Administrators' personal and social support systems*. University of Pittsburgh. (ERIC Doc. 309 526).
- Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Nias, J. (1989). *Primary teachers talking: A study of teaching as work*. London: Routledge.
- Nias, J., Southworth, G., & Yeomans, R. (1989). *Staff relationships in the primary school*. London: Cassell.
- Peters, K., & Slegers, P. (1994, april). Professionele oriëntatie van schoolleiders. *Meso*, 14(73), 22-27.
- Rosenholtz, S. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Schuit, H., Slegers, P., & Giesbers, J. (1991). *The dual character of schools as organizations*. Sheffield: Sheffield City Polytechnic.
- Smets, P. (1983). *Werk verdelen op school*. Nijmegen: ITS.
- Staessens, K. (1989). *De professionele cultuur van basisscholen in vernieuwing*. Academisch proefschrift, Katholieke Universiteit Leuven.
- Staessens, K. (1993). The professional culture at the "Cornflower". In F. Kieviet & R. Vandenbergh (Eds.), *School culture, school improvement and teacher development* (pp. 7-38). Leiden: DSWO Press.
- Stoel, W. (1994). *De taakhoud, de taakomvang en de taakbelasting van schoolleiders in het basisonderwijs*. Enschede: Universiteit Twente.
- Wit, C. de (1991). *Een beroepsprofiel voor schoolleider in het Katholiek Basisonderwijs*. 's-Hertogenbosch: KPC.

Auteurs

J. Imants (universitair docent), **C. de Brabander** (universitair docent) en **I. Bakkenes** (assistent-in-opleiding) zijn werkzaam bij de Vakgroep Onderwijsstudies van de Rijksuniversiteit Leiden.

Adres: J. Imants, Rijksuniversiteit Leiden, Vakgroep Onderwijsstudies, Postbus 9555, 2300 RB Leiden

Abstract

Principal isolation and task perception in primary schools

J. Imants, C. de Brabander & I. Bakkenes. *Pedagogische Studiën*, 1996, 73, 8-24.

Teacher isolation and teacher collaboration are regarded as important workplace conditions in schools. The concept of educational leadership stresses the central position of the principal in the school. Nevertheless, principal isolation in schools has hardly deserved any attention in educational research. The aim of this article is to develop concepts and methods for the investigation of principal isolation. Principal isolation is assumed to be a task specific concept. Isolation is determined by means of the principal's position in a communication network of the school. The research question is whether this level of participation in the network is related to principal's task perception. Results indicate that isolated principals and non-isolated principals differ in their perception of educational leadership tasks in the expected direction. Educational leadership tasks are distinguished from general management tasks and instructional tasks.

Sekse van de schooldirecteur in relatie tot de gerichtheid op het primaire proces

M.L. Krüger

Samenvatting

Het heeft tot aan de jaren zeventig geduurd voordat er in onderzoek naar schoolmanagement aandacht werd besteed aan de variabele sekse. De schaarse onderzoeksresultaten zijn bovendien nogal eens inconsistent en soms zelfs conflicterend. Dit gebrek aan wetenschappelijke kennis is aanleiding geweest voor het verrichten van een dissertatie-onderzoek naar sekseverschillen in schoolleiderschap in het Nederlandse voortgezet onderwijs. Ruim aandacht werd geschonken aan de gerichtheid van schooldirecteuren op het primaire proces; dit deel van het onderzoek zal in dit artikel centraal staan.

Met behulp van een 'matching' methode werden paren van scholen gegenereerd, steeds met een vrouwelijke en een mannelijke directeur aan het hoofd. 98 directeuren, 673 docenten en 858 leerlingen vulden een vragenlijst in. De geconstrueerde instrumenten, doorgaans Likert-schalen, werden onderworpen aan factoranalyses en betrouwbaarheidsbepalingen. De gegevens werden geanalyseerd met behulp van multivariate toetsing van gepaarde waarnemingen.

De resultaten laten zien dat de variabele 'sekse' significant van invloed is op de uitoefening van het leiderschap: vrouwen blijken sterker op het produkt van de organisatie, op het primaire proces gericht te zijn dan mannen. De resultaten blijken overigens te verschillen per schoolcultuur.

Inleiding

Het laatste decennium wordt er veel geschreven en gediscussieerd over vrouwen in de schoolleiding. Veelal wordt daarbij de nadruk gelegd op het geringe aantal vrouwelijke schoolleiders. Dit is niet zo verwonderlijk,

want ondanks het feit dat onderwijs van oorsprong een beroep is waar veel vrouwen werken, zijn de aantallen en percentages vrouwen in directeursfuncties in het onderwijs internationaal laag. Nederland spant daarin welhaast de kroon. Allerlei acties van overheid en initiatieven uit het veld ten spijt, is in ons land het aandeel van vrouwelijke directeuren in het basisonderwijs de laatste jaren slechts gestegen van ongeveer 13 naar 13,5% en in het algemeen voortgezet onderwijs van 4% naar 6%. De argumentatie voor het verhogen van het aantal vrouwen beperkt zich doorgaans dus tot het gelijke kansen principe. Argumenten die kunnen verwijzen naar de specifieke bijdragen van vrouwen in schoolleidingen, worden nauwelijks in de discussie betrokken. Nu is er ook nog nauwelijks kennis verzameld waarop een dergelijke argumentatie gestoeld zou kunnen worden. Het heeft tot de jaren zeventig geduurd voordat de variabele sekse werd betrokken in onderzoek op het gebied van schoolmanagement. Bovendien richtte dit onderzoek zich voornamelijk op loopbaanpatronen van vrouwen in onderwijsmanagement en nauwelijks op de vraag naar sekseverschillen in de wijze van uitvoering van het schoolleiderschap. Wanneer dit laatste wel het onderwerp van onderzoek was, dan richtte men zich eenzijdig op stijlen van leidinggeven. Daarbij zijn de resultaten van de verschillende onderzoeken inconsistent en soms zelfs conflicterend. Het gebrek aan wetenschappelijke kennis leidt er toe dat de maatschappelijke discussie over vrouwen in onderwijsmanagement voornamelijk is gebaseerd op vooroordelen en stereotiepen met betrekking tot schoolleiderschapskenmerken en -kwaliteiten van zowel vrouwen als mannen. Op grond van dergelijke overwegingen werd een dissertatieonderzoek verricht om na te gaan of er sekseverschillen zijn in schoolleiderschap in het Nederlandse voortgezet onderwijs, en zo ja, wat de aard van

die verschillen is (Krüger, 1994). Zowel de taken, stijlen, oriëntaties en zelfpercepties van de schooldirecteur werden onderzocht. Daarnaast werd bezien wat de effecten zijn van het leiderschap op schoolorganisatie en schoolcultuur. Naast 98 schooldirecteuren - 49 vrouwen en 49 mannen - in AVO, MBO en Dag/Avond onderwijs werden 673 docenten en 858 leerlingen bevraagd.

In het kader van het onderzoeken van taken en oriëntaties werd ruim aandacht geschonken aan het onderwijskundig leiderschap van vrouwen en mannen. De gerichtheid op het primaire proces, dus op het produkt van de organisatie, is een interessant aspect van het schoolleiderschap aangezien de rol van de schoolleider als onderwijskundig leider pregnant naar voren komt uit het schooleffectiviteits onderzoek. Deze effectiviteit van scholen staat met de autonomisering van scholen nog sterker in de belangstelling dan voorheen. In dit artikel zullen daarom alleen de resultaten van het onderzoek worden besproken die betrekking hebben op het onderwijskundig leiderschap van vrouwelijke en mannelijke schooldirecteuren.

1 Sekseverschillen in onderwijskundig leiderschap

In de onderzoeken naar de seksspecificiteit in het leidinggeven is de *stijl van leidinggeven* verreweg de frequentst gebruikte variabele, zowel in onderzoek op het terrein van bedrijfs- als van onderwijsmanagement. Doorgaans wordt deze variabele geoperationaliseerd in relaties en taakgerichtheid en soms ook in democratische versus autocratische besluitvormingsstijl. In de verschillende reviews wordt nogal eens gewezen op de methodologische zwakte van de onderzoeken (Meskin, 1974; Meskin, 1979; Fishel & Pottker, 1977; Frasher & Frasher, 1979; Adkinson, 1981). Er wordt kritiek geleverd op de meetinstrumenten; op de steekproeftrekking, -samenstelling en -omvang; op de procedures voor data-analyse; enzovoorts. De onderzoeksresultaten spreken elkaar dan ook nogal eens tegen: soms worden verschillen gevonden, soms niet. Bovendien zijn de onderzoeken op het terrein van sekse en leiderschapsstijlen in het onderwijs schaars. De 50 studies die Eagly, Karau en Johnson (1992)

aantreffen omvatten drie artikelen (Grobman & Hines, 1956; Morsink, 1970; Van Meir, 1973;), één boek (Hemphill, Griffiths & Frederiksen, 1962), één boekhoofdstuk (Charters & Jovick, 1981), 43 dissertaties en twee rapporten.

Ook in Nederland zijn eventuele verschillen in de wijze van leidinggeven tussen vrouwen en mannen nog nauwelijks onderzocht. Het onderzoek van Van de Grift en Kurek-Vriesema (1990) vormt een uitzondering op deze regel: in hun onderzoek naar het functioneren van schoolleidingen in de basisschool namen zij ook de variabele sekse op. Een andere uitzondering vormt het werk van De Jong en Van Doorne-Huiskes (1989). Hun onderzoek naar sekse en leiderschapsstijlen vond echter niet in het onderwijs plaats, maar in de overheids- en bedrijfssector. Wat de onderwijssector betreft beperkt het onderzoek naar sekse en leidinggeven in ons land zich hoofdzakelijk tot de loopbanen van vrouwen en mannen (Bonhof, De Greeuw, Pollemans & Vogelaar, 1986; Ruijs, Mens, Baggen & Jansen, 1986; Ruijs & Mientjes, 1989; Ruijs, 1990).

In onderzoeken naar schoolleiderschap definieert men doorgaans de onderwijskundig leider als iemand die gericht is op: het benadrukken van leerlingprestaties, het creëren van een op leren gericht schoolklimaat, de bemoeienis met onderwijsmethoden, het coördineren van onderwijsprogramma's, het begeleiden en ondersteunen van docenten, en het regelmatig evalueren van de voortgang van leerlingen (zie bijvoorbeeld Edmonds, 1979; Sweeney, 1982; Van de Grift, 1985b, 1985c, 1987; Van der Vegt & De Vos, 1990). Daar tegenover wordt gewoonlijk de administratieve gerichtheid van de schoolleider geplaatst: de gerichtheid op het in stand houden van de organisatie en het scheppen van randvoorwaarden voor het onderwijs. De schoolleider met deze gerichtheid houdt zich hoofdzakelijk bezig met het plannen binnen bepaalde richtlijnen, het organiseren van de randvoorwaarden van het onderwijs en het uitoefenen van controle. Een belangrijk onderscheid in taken en activiteiten van schoolleiders zou zijn de mate waarin ze het accent leggen op de primaire processen of juist de secundaire processen in de school, dat wil zeggen op onderwijs of op het scheppen

van randvoorwaarden voor het onderwijs. Schoolleiders zouden of hoofdzakelijk onderwijskundig of hoofdzakelijk administratief leiderschap uitoefenen. Onderzoeksresultaten wijzen echter uit dat schoolleiders in de praktijk met name administratieve activiteiten uitvoeren, hoewel zij als hun gewenste praktijk formuleren dat ze juist meer onderwijskundige activiteiten zouden willen uitvoeren (Gousha, 1986; Leithwood, Begley & Cousins, 1990).

Zowel in de Verenigde Staten als in Nederland werd uit eerste onderzoekingen geconcludeerd dat vrouwelijke en mannelijke schoolleiders in principe niet verschillen in de taken die ze verrichten (Shakeshaft, 1987; Ruijs, 1990). Min of meer in tegenspraak daarmee komt uit de reviews enige evidentie naar voren die wijst in de richting dat vrouwen meer activiteiten verrichten op het gebied van onderwijskundig leiderschap, terwijl mannen meer administratieve leiderschapsactiviteiten aanpakken (Meskin, 1974; Fishel & Pottker, 1977; Frasher & Frasher, 1979). De onderzoeksresultaten van Gross en Trask (1964, 1976) - de eerste onderzoekers die uitgebreid aandacht besteedden aan sekseverschillen en schoolleiderschap - bevestigen bovenstaande hypothese slechts ten dele. Zij legden aan 189 schoolleiders in het basisonderwijs, van wie bijna de helft vrouwen, lijsten voor met onderwijskundige en administratieve activiteiten. Vrouwen evalueerden hun *bekwaamheid* in het uitvoeren van onderwijskundige activiteiten significant hoger dan de mannen. Voor de schatting van de eigen bekwaamheid met betrekking tot administratieve activiteiten werd geen significant verschil gevonden, hoewel de mannen zichzelf hier iets positiever evalueerden. Vrouwen bleken een significant grotere *voldoening* te putten uit de uitoefening van onderwijskundige en mannen van administratieve activiteiten. Er werd geen significant verschil gevonden in het *belang* dat vrouwen en mannen toeschreven aan onderwijskundige activiteiten en evenmin in het belang dat men hechtte aan administratieve activiteiten. Vrouwen scoorden op beide wel iets hoger dan de mannen, op administratieve activiteiten tegengesteld aan de hypothese die Gross en Trask hadden gesteld. Gross en Trask vroegen tevens aan schoolleiders naar de criteria waarop zij hun scholen en docenten beoordelen. Dit reflecteert volgens hen de wijze

waarop schoolleiders de missie van hun school en de verwachtingen over prestaties van hun personeel definiëren. De resultaten wezen uit dat vrouwen een groter belang hechten dan mannen aan de zorg voor individuele verschillen van leerlingen, aan het helpen van leerlingen die afwijkend gedrag vertonen, en aan de zorg voor de sociale en emotionele ontwikkeling van de leerlingen. Er werd geen sekseverschil gevonden in het belang dat gehecht werd aan de twee overige factoren, namelijk leerlingprestaties en het vermogen van de school de orde te handhaven. Op een ander onderwijskundig leiderschapskenmerk blijken de seksen wel te verschillen: vrouwen oefenen significant meer controle uit over de professionele activiteiten van docenten. Mannelijke en vrouwelijke schoolleiders scoorden weer gelijk wat betreft de mate waarin zij de visie uitdragen dat de docenten maximale zorg dienen te besteden aan de leerlingen.

Charters en Jovick (1981) concludeerden uit onderzoek onder 29 schoolleiders in het basisonderwijs dat de vrouwelijke schoolleider prominenter en persoonlijker betrokken is dan haar mannelijke collega bij zaken die gerelateerd zijn aan het curriculum van de school. Shakeshaft (1987) haalt verschillende onderzoekingen aan naar de manieren waarop vrouwen en mannen het doel van hun werk conceptualiseren. Hieruit blijkt volgens haar dat vrouwen de baan van schoolleider benaderen als die van een 'master teacher or educational leader', terwijl mannen de baan bezien vanuit een 'managerial-industrial perspective'. Uit het onderzoek van Brubaker en Simon (1987) blijkt dat met betrekking tot zowel feitelijke als gewenste rollen vrouwelijke schoolleiders hoger scoren op de rol van onderwijskundig leider en mannen op de rol van 'general manager'. Ook de onderzoeksresultaten van Hallinger en Murphy (1985) en Hallinger, Bickman en Davis (1990) ondersteunen het gegeven dat vrouwen hun leiderschap meer dan mannen richten op het onderwijs.

De reden die nogal eens wordt genoemd voor het sterker op onderwijskundig leiderschap gericht zijn van vrouwen is dat ze langer doen over het doorlopen van hun carrière en dus meer ondervinding hebben voordat ze uiteindelijk schoolleider worden (Levinson, 1982; Yeakey, Johnston & Adkinson, 1986;

Hallinger et al., 1990). Wat ook de verklaring moge zijn, Angelsaksisch onderzoek geeft in elk geval aan dat vrouwen een grotere onderwijskundige en mannen een grotere administratieve gerichtheid hebben.

In Nederland hebben Van de Grift en Kurek-Vriesema (1990) onderzoek verricht in basisscholen naar de uitvoering van schoolleiderstaken, waarin zij ook de variabele sekse meenamen. Zij legden vragenlijsten voor aan 139 vrouwelijke en 168 mannelijke schoolleiders. Ook zij vonden een grotere gerichtheid van vrouwen op onderwijskundige taken, zoals het benadrukken van leerlingprestaties, het begeleiden van docenten, en het observeren in de klas. Mannen bleken meer dan vrouwen gericht op administratieve taken zoals de zorg voor het gebouw en voor de sollicitatieprocedure inclusief de opvang van nieuw personeel. Ruijs (1990) maakt in haar onderzoek naar loopbanen een uitstapje naar taakuitoefening door een lijst van 6 taakgebieden van directies voor te leggen aan schoolleiders in het basisonderwijs (309 vrouwen en 360 mannen) en in het voortgezet onderwijs (49 vrouwen en 606 mannen). Zij vindt geen sekseverschillen in portefeuilles en tijdsbesteding, ook niet wat betreft onderwijskundige en administratieve taken.

In onderzoek naar schooleffectiviteit, waarin op zoek wordt gegaan naar schoolkenmerken die de effectiviteit van scholen beïnvloeden, komt de rol van de schoolleider als onderwijskundig leider als belangrijke factor naar voren (Scheerens, 1989). In het inleidende artikel van dit themanummer wordt hier uitgebreid op ingegaan. In dit schooleffectiviteits onderzoek werd tot op heden de variabele sekse echter bij voortdurende genegeerd. Voorlopig lijken vrouwen meer gericht op onderwijskundig leiderschap dan mannen. Kunnen we nu door een en ander te verbinden, concluderen dat vrouwen effectiever zijn dan mannen? Sergiovanni is voorzichtig: er is geen harde evidentie aan te dragen, maar de literatuur over effectieve scholen laat volgens hem zien dat "while women are underrepresented in principalships, they are overrepresented in successful principalships" (Brandt, 1992, p.48).

Gezien de stand van het onderzoek is het leggen van een verband tussen sekse van de

schoolleider en schooleffectiviteit een gecompliceerde bezigheid. Bovendien is het effectieve scholen onderzoek niet gebaseerd geweest op een gemeenschappelijk theoretisch model, waardoor eenduidige conceptualiseringen van de termen schooleffectiviteit en onderwijskundig leiderschap ontbreken. Het verband tussen de schoolleiderssekte en schooleffectiviteit berust dus nog voor het grootste deel op speculaties.

2 Onderzoeksvraagstelling en hypothesen

Zoals uit het voorgaande blijkt is er weinig onderzoek verricht naar sekse en leiderschap en nog minder naar sekse en onderwijskundig leiderschap, ondanks het feit dat onderwijskundig leiderschap steeds verbonden wordt met de effectiviteit van scholen. Het weinige onderzoek dat werd verricht naar sekse en onderwijskundig leiderschap richtte zich bijna uitsluitend op het basisonderwijs. Vanwege de andere taakstelling van schoolleiders in het basisonderwijs kunnen de resultaten niet zonder meer worden overgezet naar het voortgezet onderwijs. Bovendien zijn er overwegingen met betrekking tot de onderwijskundige gerichtheid van schoolleiders in Nederland, die tot voorzichtigheid manen bij het geldig verklaren van de Angelsaksische onderzoeksresultaten voor de Nederlandse situatie. Weliswaar hebben ook in ons land vrouwen een langere loopbaan als docent voordat ze in een leidinggevende positie terecht komen, hetgeen zoals hiervóór bleek een sterke onderwijskundige gerichtheid tot gevolg zou kunnen hebben. Aan de andere kant zijn er echter factoren in de context van het leidinggeven in ons land die ertoe zouden kunnen leiden dat schoolleiders, of ze nu man of vrouw zijn, minder gericht zijn op de onderwijskundige kant van het leidinggeven dan elders. Zo is er bijvoorbeeld geen diplomering van schoolleiders vereist. Schooldirecteuren in functie zijn dus niet noodzakelijkerwijs geschoold in het managen van een onderwijsinstelling. Schaalvergroting en decentralisatie vergroten daarnaast de noodzaak tot het leggen van nadruk op de organisatorische kant van het managen van een school; ze vergen eerder administratief dan onderwijskundig leiderschap.

Misschien samenhangend hiermee verschaft in Nederland de bemoeienis met onderwijszaken de schoolleider een lage status; de schoolleider die organisatiezaken en externe contacten in portefeuille heeft, wordt als manager gewaardeerd. Omdat hier de status van de leider in het geding is, zou op dit punt de druk van de organisatie tot socialisatie in de leiderschapsrol (Eagly & Johnson, 1990) voor beide seksen wel eens groot kunnen zijn. Vrouwen en mannen zullen dan evenzeer geneigd zijn om zich meer bezig te houden met administratief dan met onderwijskundig leiderschap. Sekseverschillen worden daardoor voor een deel genivelleerd. Een van de vraagstellingen uit mijn dissertatieonderzoek die in dit artikel centraal wordt gesteld, luidt daarom: *zijn er verschillen tussen vrouwelijke en mannelijke schooldirecteuren in het voortgezet onderwijs in Nederland met betrekking tot hun onderwijskundige en administratieve taken en oriëntaties?*

Wat betreft de *taken* van schooldirecteuren werd onderzoek gedaan naar portefeuilles, tijdsbesteding en naar de mate waarin belang wordt gehecht aan onderwijskundige en administratieve taken. Hun *oriëntatie* werd op handeliningsniveau onderzocht door de directeuren en hun docenten te vragen naar de activiteiten die directeuren verrichten op onderwijskundig terrein. Vier van de 22 hypothesen die voor het onderzoek werden geformuleerd zijn van toepassing op de in dit artikel behandelde vraagstelling:

Hypothese 1 Er is geen verschil in de takenpakketten van vrouwen en mannen als schooldirecteur.

Hypothese 2 Vrouwen besteden in hun functie van schooldirecteur meer tijd aan onderwijskundige taken en mannen meer aan administratie en beheer.

Hypothese 3 In tegenstelling tot mannen hechten vrouwen in hun directeursfunctie meer belang aan onderwijskundige taken dan aan administratieve taken.

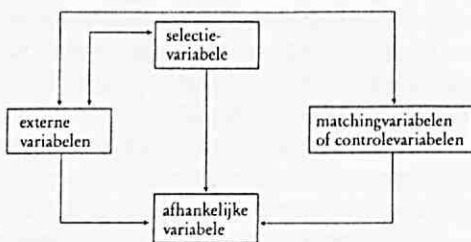
Hypothese 4 Vrouwen zijn in hun handelingen als schooldirecteur meer gericht op onderwijskundig leiderschap dan mannen.

3 Onderzoeksoptzet en uitvoering

Op grond van de vraagstelling werd besloten tot een empirisch vergelijkend onderzoek onder schooldirecteuren in het voortgezet onderwijs: AVO, MBO en Dag/Avond onderwijs. Het 'General Framework of Instructional Leadership', voortgekomen uit onderzoek van het Far West Laboratory (Bossert, Dwyer, Rowan & Lee, 1982; Dwyer et al., 1985; Dwyer, Barnett & Lee, 1987) vormde het fundament voor de constructie van een eigen onderzoeksmodel. Met behulp van een 'matchings'procedure werden paren van bij elkaar passende scholen gezocht, waarbij de ene school door een vrouw en de andere school door een man werd geleid. Daarvoor werden eerst alle vrouwelijke directeuren in het Nederlands voortgezet onderwijs die konden worden opgespoord aangeschreven. Meer dan 80% van hen reageerde positief. Na de 'matching' met scholen met mannelijke directeuren en na het afvallen van een aantal directeuren met name vanwege scholenfusies bestond de populatie uiteindelijk uit 98 schooldirecteuren: 49 paren van scholen met steeds een vrouw en een man aan het hoofd. Elke school werd bezocht door een doctoraalstudent onderwijskunde. Zij kregen opdracht om docenten en leerlingen random te selecteren uit de schooladministratie bestanden. Op basis van een power-analyse (Cohen, 1988) werd bepaald dat er - rekening houdend met een alpha van .05 - per school 5,9 observaties nodig zouden zijn om een power van minimaal .90 te bereiken voor een gemiddelde effectgrootte. Rekening houdend met een eventuele non-respons werden de aantallen bepaald op 10 leerlingen en 9 docenten per school. De docenten moesten gelijk worden verdeeld over de seksen, gelijk worden gespreid over alfa-, beta- en gammavakken en mochten niet na 1 augustus 1989 zijn aangesteld. Bij de selectie van de leerlingen moest worden gezorgd voor een gelijke spreiding over de seksen en over de leerjaren, waarbij geen leerlingen uit het eerste leerjaar werden geselecteerd. Directeuren, leerlingen en docenten kregen een vragenlijst voorgelegd. Alle directeuren, meer dan 90% van de leerlingen (858) en meer dan 80% van de docenten (673) vulden hun vragenlijst in. Ten behoeve van de 'matching' werd ge-

bruik gemaakt van de 'guaranteed variable caliper matching' van Althausen en Rubin (1970). Met deze individuele 'matching'-methode wordt voor elke persoon in de ene groep met behulp van een computerprogramma gezocht naar een persoon met zo gelijk mogelijke kenmerken in de andere groep. Eerst wordt naar perfecte 'matches' gezocht. Wanneer er niet voldoende perfecte 'matches' kunnen worden gevonden, worden de marges automatisch vergroot. De kwaliteit van de 'matches' wordt op deze manier zo hoog mogelijk gehouden. 'Matching' stelt de onderzoeker in staat om variabelen waarvan bekend is of verondersteld kan worden dat ze van invloed zijn op de afhankelijke variabele, in dit onderzoek de uitoefening van het schoolleiderschap, onder controle te houden. De invloed van een aantal storende variabelen wordt zo bij voorbaat geneutraliseerd. Op die manier kon de 'netto-invloed' zoals Peschar (1975) en Bosman en Louwes (1982) dat noemen, van de factor sekse worden vastgesteld op de wijze van uitoefening van het schoolleiderschap. Vanwege het lage percentage vrouwelijke schooldirecteuren in Nederland ten opzichte van het percentage mannen leek 'matching' een zeer geschikte methode voor dit specifieke onderzoek. Meester en De Leeuw (1983) achten paren matching misschien wel de beste methode om een klein deel van het oorspronkelijke bestand te selecteren.

Althausen en Rubin (1970, pp.326-328) onderscheiden naast de afhankelijke variabele drie soorten onafhankelijke variabelen, namelijk de selectievariabele, de controlevariabelen en de externe variabelen. In *Figuur 1* zijn de typen variabelen in schema gezet.



Figuur 1. Typen variabelen bij matching

De *selectievariabele* is de onafhankelijke variabele waarop onderscheid tussen de populaties wordt aangebracht, in dit geval sekse.

Daarnaast zijn er twee andere typen onafhankelijke variabelen, dat wil zeggen variabelen die verband kunnen houden met of van invloed kunnen zijn op de afhankelijke variabele en waarvan dus te verwachten is dat ze een eventueel verband tussen de selectievariabele en de afhankelijke variabele zullen vertekenen of verdoezelen. Wanneer deze variabelen door middel van 'matching' worden gecontroleerd noemen we ze '*matching*'variabelen. In het hier besproken onderzoek werden 11 context factoren gekozen als 'matching'variabele, namelijk schooltype; schoolgrootte; denominatie; urbanisatie; seksesamenstelling directieteam, docenten- en leerlingenpopulatie; leeftijd docentenpopulatie; percentage allochtone leerlingen; leeftijd schooldirecteur; lengte van het directeurschap. Alleen vanwege de gunstige numerieke verdeling van de populaties kon op dit grote aantal variabelen paarsgewijs worden gelijkgeschakeld. Wanneer dergelijke variabelen niet onder controle worden gehouden, heten ze *externe variabelen*. Redenen om de groepen of individuen op bepaalde variabelen niet gelijk te schakelen kunnen zijn, dat gegevens over deze variabelen niet beschikbaar zijn, dat de invloed ervan onbekend is of dat het aantal 'matching'variabelen dient te worden beperkt. Als besloten wordt om een onafhankelijke variabele niet in de 'matching'-procedure op te nemen, dan wordt het dus direct een externe variabele. Voorbeelden van dergelijke variabelen in dit onderzoek zijn: vooropleiding directeur, vol- en deeltijd werken docenten, SES leerlingeninstroom, dislokatie.

De data van de 98 schooldirecteuren werden paarsgewijs aaneengeschaakt tot 49 gepaarde waarnemingen. De data van de docenten en leerlingen werden geaggregeerd om ze vervolgens naar het niveau van de directeuren te brengen; ze werden dus aangehaakt aan het directeurenbestand. Alle gegevens werden vervolgens op schoolleidersniveau gemeten, hetgeen mogelijk was omdat er geen hypothesen waren geformuleerd met betrekking tot de interactie van schoolleiders- en docentvariabelen. Multiniveau-analyse was dan ook overbodig. Voor de toetsing van de hypothesen werd dientengevolge gebruik gemaakt van variantieanalyse voor herhaalde metingen (SPSS MANOVA en ANOVA). Als significantien-

veau werd een alpha van .05 aangehouden. De 'matching' variabelen en de externe variabelen die in de t-toets voor gepaarde waarnemingen significante verschillen vertoonden, werden in de analyses betrokken als covariaten. Variabelen die een significante correlatie lieten zien met een afhankelijke variabele die op dat moment werd geanalyseerd, werden eveneens als covariaat in de analyse meegenomen.

4 Instrumentconstructie

De instrumenten die werden geconstrueerd ter meting van het onderwijskundig leiderschap worden in deze paragraaf kort besproken. Alle instrumenten werden eerst voorgelegd aan beoordelaars. Voor een uitgebreidere bespreking van de instrumentconstructie wordt verwezen naar Krüger (1994).

Met behulp van een een lijst van 8 taken voor schooldirecteuren werd onderzocht welke taken men zelf in portefeuille heeft respectievelijk delegeert, en wat de rangordening in taken is naar de hoeveelheid tijd die men aan elke taak besteedt. De lijst van taken ontstond uit een combinatie van de takenlijsten van Ruijs (1990) en Mahieu (z.j.) (zie Tabel 1 en 2). Voorbeeldomschrijvingen werden toegevoegd.

Om te onderzoeken welk belang vrouwen en mannen hechten aan de uitoefening van onderwijskundige en administratieve taken werden twee schalen van Gross en Trask (1976) gecombineerd. De schalen bevragen respectievelijk het belang dat men toekent aan 'Supervision of instruction' (8 items) en 'Administrative tasks' (11 items). De items werden vertaald en zo nodig aangepast aan de Nederlandse situatie. Voorbeelden van onderwijskundige items zijn 'Introduceren van nieuwe onderwijsideeën', 'Zorg dragen voor nascholingsprogramma's van docenten'. Administratieve items zijn onder andere 'beheren van de financiën' en 'roosteren'. De 18 oorspronkelijke items, voorzien van zes antwoordmodaliteiten, werden dooreen gemengd aangeboden als één instrument.

Factoranalyse wees uit dat het instrument de twee factoren onderwijskundig en administratief leiderschap goed onderscheidt. Op grond

van factor- en item-analyse werden drie items verwijderd. Het instrument 'Belang hechten aan het uitoefenen van onderwijskundige en administratieve taken' omvat uiteindelijk twee Likertschalen:

- *Belang hechten aan het uitoefenen van administratieve taken* (8 items): Cronbachs α .81
- *Belang hechten aan het uitoefenen van onderwijskundige taken* (7 items): Cronbachs α .70

De onderwijskundige gerichtheid van schoolleiders werd onderzocht als de mate waarin schooldirecteuren zeggen activiteiten te verrichten op onderwijskundig terrein en de mate waarin hun docenten deze activiteiten percipiëren. Er werd voor een operationalisatie op handelingsniveau gekozen, omdat er bij bevraging op attitudeniveau een gereede kans zou bestaan op sociaal wenselijke beantwoording. Bijna elke schoolleider realiseert zich inmiddels immers, dat onderwijskundig leiderschap een belangrijke factor is in de verhoging van de schooleffectiviteit. Het voor dit doel geconstrueerde instrument '*Onderwijskundige Gerichtheid van Schoolleiders*' (OGS) vindt zijn oorsprong in twee bestaande instrumenten, namelijk van Hallinger (Hallinger et al., 1983; Hallinger & Murphy, 1987; Hallinger, 1989a, 1989b) en van Van de Grift (1985a, 1985b).

Hallinger en collega's ontwikkelden een model van onderwijskundig leiderschap dat zou moeten kunnen verklaren hoe scholen via het leiderschap van de leider effectief worden. Zij operationaliseerden het begrip onderwijskundig leiderschap in drie dimensies:

- Definieren van een schoolmissie
- Leiding geven aan het onderwijsprogramma
- Bevorderen van een op leren gericht klimaat met in totaal 10 indicatoren die schoolleidersfuncties vertegenwoordigen, verdeeld over de 3 dimensies. Op grond daarvan ontwierp Hallinger voor elke schoolleidersfunctie een schaal, dus tien schalen in totaal. Per schaal formuleerde hij 5 items als operationalisatie van gedragingen van schoolleiders die bij een functie behoren. Zo ontstond de '*Principal Instructional Management Rating Scale*' (PIMRS). Het instrument dat ontworpen werd voor het voortgezet onderwijs bevat dus 50 specifieke

leiderschapsgedragingen. De PIMRS voor schoolleiders werd later bewerkt voor docenten en kent dus een 'principal form' en een 'teacher form'.

Daarnaast werd één van de instrumenten van Van de Grift gebruikt, namelijk de vragenlijst 'Onderwijskundig leiderschap'. De lijst is ontworpen voor schoolleiders in het basisonderwijs, hetgeen betekende dat de items moesten worden bewerkt voor het voortgezet onderwijs.

Voor de constructie van de OGS werden de dimensies en schalen van Hallinger als uitgangspunt genomen; geschikte schalen en items uit het instrument van Van de Grift werden toegevoegd. De OGS bestaat uit 10 schalen verdeeld over 3 dimensies, met in totaal 52 items. Het instrument werd tevens bewerkt ter onderzoeking van de docentperceptie van de onderwijskundige gerichtheid van hun school-directeur. Aan schooldirecteuren wordt gevraagd om met behulp van zes antwoordmodaliteiten aan te geven hoe vaak men de 52 onderwijskundige activiteiten verricht; aan docenten om met behulp van dezelfde zes antwoordmodaliteiten aan te geven in welke mate naar hun waarneming hun schooldirecteur deze 52 activiteiten verricht.

Orthogonaal noch oblique verrichte factoranalyses wezen de oorspronkelijke structuur van de OGS aan: schalen noch dimensies werden onderscheiden. Wel werd in zowel de docentversie als de schoolleidersversie van de OGS een andere, inhoudelijk heel zinvolle, structuur zichtbaar. De twee versies van de OGS bleken elkaar dus te valideren. Op grond van de factoranalyses werd besloten tot de hieronder weergegeven nieuwe instrumentstructuur. In verband met lage interitemcorrelaties werd achteraf uit 2 schalen 1 item verwijderd. Het uiteindelijke instrument bestaat daarmee uit 6 schalen met 50 items. De schalen moeten worden opgevat als zes dimensies van onderwijskundig leiderschap.

Schaal I 'Missiegericht onderwijskundig leiderschap' (14 items)

docenten α .91; directeuren α .85

Subschaal 'Het stellen van schooldoelen' (5 items)

docenten α .77; directeuren α .66

Subschaal 'Het uitdragen van schooldoelen' (5 items)

docenten α .80; directeuren α .76

Subschaal 'Het coördineren van het curriculum' (4 items)

docenten α .73; directeuren α .59

Schaal II 'Cultuurgericht onderwijskundig leiderschap' (14 items)

docenten α .92; directeuren α .82

Subschaal 'Het geven van individuele begeleiding aan docenten' (5 items)

docenten α .81; directeuren α .66

Subschaal 'Het streven naar samenwerking en consensus in het team' (5 items)

docenten α .88; directeuren α .77

Subschaal 'Het zichtbaar zijn in de school' (3 items)

docenten α .75; directeuren α .64

Schaal III 'Klassegericht onderwijskundig leiderschap' (7 items)

docenten α .78; directeuren α .68

Subschaal 'Het begeleiden en evalueren van onderwijs' (5 items)

docenten α .77; directeuren α .67

Daaraan toegevoegd: 2 items uit een andere oorspronkelijke schaal

Schaal IV 'Onderwijskundig leiderschap via het betonen van docenten' (5 items)

docenten α .75; directeuren α .56

Schaal V 'Onderwijskundig leiderschap via het bevorderen van nascholing' (5 items)

docenten α .85; directeuren α .75

Schaal VI 'Onderwijskundig leiderschap via het werken aan een ordelijke en taakgerichte sfeer' (6 items)

docenten α .82; directeuren α .78

De betrouwbaarheidscoëfficiënten van de schalen en subschalen in de docentversie van de OGS zijn hoog: de α 's liggen ruim tot zeer ruim boven de vereiste .70 (Nunnally, 1978, p. 245). Hoewel iets lager dan in de docentversie is ook de betrouwbaarheid van de schalen in de schoolleidersversie prima, met uitzondering van schaal 4. De α 's van de subschalen halen niet altijd de .70. Hierbij moet in het oog worden gehouden dat er voor de schoolleidersversie van de OGS beduidend minder proefpersonen beschikbaar zijn om het instrument te toetsen dan voor de docentversie (98 versus 673).

5 Onderzoeksresultaten: Sekse van de schooldirecteur en de gerichtheid op het primaire proces

5.1 Onderwijskundige en administratieve portefeuilles en tijdsbesteding

De in paragraaf 4 besproken lijst met 8 taakgebieden werd allereerst aan de schooldirecteuren voorgelegd met de vraag om voor elke taak aan te geven of men deze zelf in portefeuille heeft, samen met een ander beheert, of geheel delegeert. De resultaten van de multivariate analyse worden weergegeven in Tabel 1.

Zowel wat betreft *onderwijskundige zaken* als *administratie en beheer*, de taken waarnaar in dit artikel de interesse uitgaat, doen zich sekseverschillen in portefeuilles voor. Vrouwen behartigen de portefeuille onderwijskundige zaken vaker zelf dan hun mannelijke collega's. In eerste instantie lijken mannen vaker dan vrouwen administratie en beheer zelf in portefeuille te hebben. Na analyse met invoering van covariaten (zie Tabel 1) blijkt dit verschil ech-

ter op zijn minst te moeten worden genuanceerd. In de situatie dat directeuren werden benoemd in een heterogene cultuur blijken mannen administratie en beheer vaker zelf in portefeuille te houden dan vrouwen. Bij directeuren die werden benoemd in homogene en meerstromige culturen worden geen sekseverschillen gevonden (zie voor cultuurbeschrijvingen Marx, 1986). Daarbij moet worden aangetekend dat de aantallen directeuren benoemd in een homogene of in een meerstromige cultuur erg klein zijn, respectievelijk 8 en 7 directeuren.

Bovendien hebben mannen administratie en beheer vaker in portefeuille dan onderwijskundige zaken ($p = .001$), terwijl vrouwen beide portefeuilles nagenoeg even vaak zelf behartigen ($p = ns$).

Vervolgens werd dezelfde lijst met taken opnieuw voorgelegd nu met de vraag om de taken in volgorde van tijdsbesteding te plaatsen, onafhankelijk van of men de taak al dan niet zelf in portefeuille heeft. Als gevolg van het feit dat de directeuren een rangordening hebben gegeven van 1 tot 8 is het antwoord op de laatste taak per respondent voorspelbaar; daardoor is

Tabel 1
Portefeuilles (1 = zelf in portefeuille; 2 = zelf en ander; 3 = gedelegeerd)

	V	M	p
Organisatie ontwikkeling	1,20	1,12	ns
Interne communicatie	1,53	1,63	ns
Onderwijskundige zaken	1,67	2,22	effect sekse .002** effect schoolsoort .002 interactie-effect sekse en schoolsoort ns
Personeelsbeheer	1,20	1,27	ns
Leerling/stagebegeleiding	2,74	2,92	.05* → cov. schoolgrootte .07; sekseverschil ns
Administratie & beheer	1,94	1,55	.04* → cov. cultuur entree .002; sekseverschil ns
Relatie met buitenwereld	1,18	1,29	ns
Marketing	1,80	2,00	ns

Tabel 2
Tijdsbesteding aan taken (1 = minste tijd; 8 = meeste tijd)

	V	M	p
Organisatie ontwikkeling	6,4	6,0	ns
Interne communicatie	5,4	4,6	.04* → cov. sekse leerlingen .03; sekseverschil .21 → cov. mann. docenten .04; sekseverschil .13
Onderwijskundige zaken	5,5	4,6	effect sekse .06* effect schoolsoort .008** interactie-effect sekse en schoolsoort ns
Personeelsbeheer	5,6	5,0	ns
Leerling/stagebegeleiding	2,1	2,0	ns
Administratie & beheer	3,5	4,5	.007**
Relatie met buitenwereld	4,6	5,8	.001**
Marketing	3,0	3,5	ns

Tabel 3*De portefeuille Onderwijskundige zaken en de tijdsbesteding daaraan uitgesplitst per schoolsoort*

	Portefeuille (1=zelf; 2=zelf/ander; 3=gedelegeerd)			Tijdsbesteding (1 = minste tijd; 8 = meeste tijd)		
	V	M	p	V	M	p
AVO (26)	1,46	1,92	.05*	6,3	4,8	.02*
MBO (13)	2,23	2,54	ns	4,1	3,9	ns
D/A (10)	1,50	2,60	.02*	5,5	4,9	ns
alle scholen (49)	1,67	2,22	.002**	5,5	4,6	.05*

een multivariate analyse van de antwoorden niet mogelijk. Daarom worden in Tabel 2 de resultaten van de univariate analyses per taak weergegeven.

Ook wat de tijdsbesteding betreft doen zich sekseverschillen voor in de onderwijskundige en de administratieve taken. Vrouwen verschillen echter ook in hun tijdsbesteding aan relaties met de buitenwereld, hetgeen in het verband van dit artikel eveneens een interessant sekseverschil is. Op de overige taken verschillen de seksen niet in tijdsbesteding. Vrouwelijke directeuren besteden meer tijd aan onderwijskundige zaken dan mannen, die op hun beurt meer tijd besteden dan vrouwen aan administratie en beheer en aan het onderhouden van relaties met de buitenwereld. Bovendien besteden vrouwen meer tijd aan onderwijskundige zaken dan aan administratie en beheer ($p = .000$), terwijl mannen aan deze twee taken evenveel tijd besteden ($p = ns$). Aan externe relaties besteden mannen meer tijd dan aan onderwijskundige zaken ($p = .003$), terwijl dat bij vrouwen andersom is ($p = .045$).

Zowel bij analyse van de portefeuille onderwijskundige zaken als bij analyse van de tijdsbesteding daaraan bleken de seksen verschillend te antwoorden voor de onderscheiden schoolsoorten (effect schoolsoort in beide gevallen .008). Voor administratie en beheer en relatie met de buitenwereld bleek dit niet het geval te zijn. Daarom worden voor onderwijs-

kundige zaken ook de verschillen tussen de schoolsoorten geanalyseerd (zie Tabel 3).

De tabel maakt zichtbaar dat vrouwen alleen in het AVO onderwijskundige zaken vaker in portefeuille hebben en er ook meer tijd aan besteden. In het Dag/Avond onderwijs hebben zij de taak wel vaker zelf in portefeuille, maar de tijdsbesteding eraan verschilt niet significant. In het MBO blijkt noch de portefeuille noch de tijdsbesteding te verschillen voor vrouwen en mannen.

Er is dus wel degelijk verschil in takenpakketten van vrouwen en mannen, waarmee *hypothese 1* in zijn algemeenheid dient te worden verworpen. Vrouwen hebben – in elk geval in het AVO en het Dag/Avond onderwijs – vaker dan mannen onderwijskundige zaken zelf in portefeuille. In het geval dat directeuren benoemd werden in heterogene culturen hebben mannen vaker dan vrouwen de taak administratie en beheer in eigen portefeuille.

Hypothese 2 wordt aanvaard: vrouwen besteden meer tijd aan onderwijskundige zaken en mannen meer aan administratie en beheer en aan de relatie met de buitenwereld.

5.2 Het belang hechten aan onderwijskundige en administratieve taken

Ter toetsing van *hypothese 3* wordt gekeken naar sekseverschillen in het belang dat men hecht aan onderwijskundige en aan administratieve taken (zie Tabel 4).

Tabel 4*Het hechten van belang aan onderwijskundige en administratieve taken (1 = geen enkel belang; 6 = extreem belangrijk)*

	V	M	p sekseverschil
Onderwijskundige taken	4,1	3,6	.000**
Administratieve taken	3,1	3,0	ns
p binnen de seksen	.000**	.000**	

Zowel vrouwelijke als mannelijke directeuren leggen meer accent op het primaire dan op het secundaire proces, waarmee *hypothese 3* wordt verworpen. Wanneer Tabel 4 horizontaal wordt geïnterpreteerd, komt er echter toch een interessant sekseverschil naar voren: vrouwen hechten meer belang aan onderwijskundige activiteiten dan mannen. Aan administratieve activiteiten hechten de seksen evenveel belang. Dit geldt zowel voor het AVO, het MBO als het Dag/Avond onderwijs (het effect van het onderscheid in schoolsoorten en het interactie-effect schoolsoort-sekse zijn voor onderwijskundige noch voor administratieve taken significant). Opname van de significante covariaten - 'seksleerlingen' en 'cultuur bij benoeming' in geval van administratieve activiteiten; 'wel/geen kinderen hebben' en 'aantal vrouwelijke docenten' in geval van onderwijskundige activiteiten - leidt niet tot verandering van de resultaten.

5.3 Het verrichten van activiteiten op het gebied van onderwijskundig leiderschap

In het voorgaande werd met betrekking tot het onderwijskundig leiderschap onderzocht of men de taak zelf in portefeuille heeft, de tijd die

men eraan besteedt en het belang dat men eraan hecht. In deze paragraaf wordt de gerichtheid op de primaire processen vertaald als de mate waarin schooldirecteuren - volgens henzelf en volgens hun docenten - onderwijskundige activiteiten verrichten.

Allereerst werd een multivariate analyse uitgevoerd op de zes hoofddimensionen van de OGS, waarbij onderscheid werd gemaakt in drie schoolsoorten, maar nog geen rekening werd gehouden met covariaten. Daarna zijn per dimensie de significanties univariaat berekend, waarbij opnieuw onderscheid in schoolsoorten werd gemaakt en waarbij ook de significante covariaten werden betrokken. Wanneer de scores van de directeuren niet bleken te verschillen voor de onderscheiden schoolsoorten en er evenmin sprake was van een sekseverschil in antwoordpatronen in de drie schoolsoorten, werd de analyse opnieuw uitgevoerd zonder onderscheid in schoolsoorten te maken. De resultaten van de multivariate en de univariate analyses voor directeuren en voor docenten worden tezamen in Tabel 5 weergegeven.

Een eerste blik op de tabel laat direct al duidelijk zien dat vrouwen een sterkere onderwijs-

Tabel 5

Onderwijskundige Gerichtheid van Schoolleiders volgens schooldirecteuren zelf en volgens docenten (1 = nooit; 6 = heel vaak)

	Directeuren			Docenten		
	V	M	p	V	M	p
Missiegerichtheid	4,0 → cov. sekse docs. .04: p=.036*	3,6	.009**	3,5	3,3	.007**
Cultuurgerichtheid	4,5 → covs. sekse docenten en urbanisatie .01: p=.010**	4,1	.014*	3,7 → covs. sekse docs., dislokatie en schoolgrootte .02: p=.037*	3,4	.002**
Klassegerichtheid	2,9 → covs. sekse docs. en wel/geen lestaak directeur .03: p=ns	2,8	ns	2,4 → cov. schoolgrootte .04: p=.050*	2,1	.007**
Belonen docenten	3,1	2,6	.001**	2,5	2,1	.001**
Stimuleren nascholing	4,2	3,7	.002**	3,3	2,8	.000**
Taakgerichte sfeer	3,7	3,6	ns	3,4 → covs. sekse docenten en leeftijd adjuncten .001: p=ns	3,3	ns
effect sekse:	p = .007**			p = .000**		
effect schoolsoort:	p = .001**			p = .001**		
interactie-effect sekse en schoolsoort:	p = ns			p = .000**		

kundige gerichtheid hebben dan mannen. Volgens de directeuren is dit het geval voor 4 van de 6 en volgens de docenten zelfs voor 5 van de 6 dimensies. Opvallend is dat de seksen juist wat betreft de dimensies die het dichtst bij de leerling staan, de dimensies 'Klassegerichtheid' en 'Werken aan een ordelijke en taakgerichte sfeer', volgens de directeuren zelf niet verschillen. Naar de perceptie van de docenten scoren vrouwen overigens wel hoger op 'Klassegerichtheid' dan mannen.

Hoewel niet direct relevant in het kader van sekseverschillen is het toch interessant dat Tabel 5 laat zien, dat directeuren zichzelf op alle dimensies hoger inschatten dan hun docenten. Redenen hiervoor kunnen zijn dat schooldirecteuren zichzelf overschatten of dat zij de frequentie waarin ze onderwijskundige activiteiten verrichten onopzettelijk overdrijven om redenen van sociale wenselijkheid. Wanneer dit het geval is, geeft de perceptie van hun docenten een objectiever beeld van hun activiteiten. Het is echter ook mogelijk dat de verschillende scores veroorzaakt worden doordat de directeuren beter dan hun docenten op de hoogte zijn van welke activiteiten ze verrichten; docenten zullen immers slechts een be-

paald percentage van de activiteiten van hun directeur waarnemen.

Alle schalen bijeengenomen, blijken de antwoorden van de directeuren te verschillen in AVO, MBO en Dag/Avond onderwijs ($p = .001$), maar mannen en vrouwen hebben blijkbaar eenzelfde antwoordpatroon in de drie onderscheiden schoolsoorten ($p = ns$). Bij de docenten ligt dit anders: ook zij antwoorden verschillend per schoolsoort ($p = .001$), maar docenten verschillen in de drie onderscheiden schoolsoorten óók wat betreft hun perceptie van sekseverschillen in onderwijskundige activiteiten van directeuren ($p = .000$). Om de richting van deze verschillen te onderzoeken werd elke schaal en elke subschaal afzonderlijk geanalyseerd op verschillen tussen de schoolsoorten, waarbij ook de covariaten in de beschouwing werden betrokken. De resultaten van deze analyses worden voor de directeuren en de docenten afzonderlijk samengevat in Tabel 6.

De scores van de directeuren- en de docentenpopulatie komen sterk overeen. De docentenscores valideren dus evenals de factoranalyses (par.4) de scores van de directeuren. Dit wijst

Tabel 6

Samenvatting resultaten Onderwijskundige Gerichtheid van Schoolleiders volgens directeuren en docenten

	Alle scholen		AVO		D/A	MBO
Missiegerichtheid	dir*	doc**	dir*	doc*		doc**
- stellen doelen	dir*	doc*	dir*			doc**
- uitdragen doelen	dir*	doc*	dir*			doc**
- coördineren curriculum		doc**	dir*	doc**		doc**
						geen
Cultuurgerichtheid	dir*	doc*	dir*		dir*	doc**
- begeleiding docenten	dir**	doc*	dir**	doc*	dir+	doc**
- samenwerking team						doc**
- zichtbaar in school	dir+				dir*	doc**
						seks-
Klassegerichtheid		doc*		doc**		doc**
Belonen docenten	dir**	doc**	dir**	doc**	dir+	doc**
						verschillen
Stimuleren nascholing	dir**	doc**	dir*	doc**	dir+	doc**
Taakgerichte sfeer						doc**

dir : volgens directeuren zijn vrouwen onderwijskundiger gericht dan mannen

doc : volgens docenten zijn vrouwen onderwijskundiger gericht dan mannen

** : $p \leq .01$

* : $p \leq .05$

+ : $p = \text{trend } (>.05 \text{ en } <.08)$

erop dat de OGS zodanig is geconstrueerd dat er inderdaad één achterliggend construct per schaal aanwezig is.

Het mag een opvallend resultaat worden genoemd, dat op geen van de schalen of sub-schalen mannen significant hoger scoren dan vrouwen; of vrouwen verrichten meer onderwijskundige activiteiten dan mannen, of er wordt geen sekseverschil gevonden.

Een ander saillant resultaat is dat in het MBO bij directeuren én docenten op geen van de zes dimensies een sekseverschil in onderwijskundige gerichtheid wordt aangetroffen. In het AVO worden, gelet op de antwoorden van de directeuren, op vier van de zes dimensies door vrouwen meer activiteiten uitgevoerd dan door mannen en in het Dag/Avond onderwijs op drie van de zes. Volgens de docenten geldt voor het AVO dat op vier van de zes dimensies vrouwen hoger scoren, terwijl in het Dag/Avond onderwijs volgens hen het sekseverschil zelfs op alle zes de schalen en op alle subschalen hogelijk significant is, ondanks het kleine aantal cases.

Daarmee is *hypothese 4* bevestigd voor het AVO en het Dag/Avond onderwijs, en verworpen voor het MBO: vrouwen in het AVO en het Dag/Avond onderwijs zijn in hun handelingen meer gericht op onderwijskundig leiderschap dan mannen.

6 Discussie

'Matching' bleek een geschikte methode voor het onderzoeken van sekseverschillen in schoolleiderschap, met name omdat met een 'matchings' procedure gemakkelijk de contextfactoren in onderzoek naar schoolleiderschap kunnen worden betrokken. Beperken we ons tot het Nederlandse onderzoek, dan zien we dat bijvoorbeeld Ruijs (1990) noch Van de Grift en Kurek-Vriesema (1990) achterliggende factoren analyseren die althans een gedeelte van de verklaring zouden kunnen vormen voor het wel of juist niet vinden van sekseverschillen.

Anderzijds kon de 'matchings' procedure geen soelaas bieden voor het probleem van selectie van misschien een speciaal soort vrouwen: met zo weinig vrouwen in schoolleidingen is immers de kans groot dat een specifiek type vrouwen is benoemd. Bovendien zijn

scholen die een vrouw hebben benoemd misschien ook wel speciale scholen. De conclusie hieruit zou kunnen zijn dat de sekseverschillen die werden gevonden mede veroorzaakt worden doordat de onderzochte vrouwen gemiddeld beter zijn dan de vrouwen die nog benoemd gaan worden. Dit mechanisme zou de sekseverschillen vergroot kunnen hebben. De vrouwengroep is dus wel representatief voor vrouwelijke schooldirecteuren op dit moment, maar misschien niet voor de toekomstige groep vrouwelijke directeuren.

Een andere wellicht even geldige redenering die hier haaks op staat, is dat de nu benoemde vrouwen meer 'mannelijk gedrag' vertonen dan ze zouden willen, omdat men nog niet erg gewend is aan vrouwelijk leiderschap en vrouwen zich moeten handhaven in een mannenwereld. Misschien zijn ze mede daarom geselecteerd. De conclusie hieruit zou zijn dat de gevonden sekseverschillen door selectiemechanismen en socialisatie in de organisatie juist genivelleerd zijn. Er is een tweede argument voor het vermoeden dat de sekseverschillen die in dit onderzoek naar voren zijn gekomen eerder kleiner dan groter zijn dan in de populaties van vrouwelijke en mannelijke schooldirecteuren. Dit argument komt voort uit de mate van representativiteit van de populatie mannelijke directeuren. De schoolkenmerken van de geselecteerde mannen komen hoogstwaarschijnlijk sterker overeen met die van vrouwelijke directeuren dan met de schoolkenmerken van de gemiddelde mannelijke directeur. De geselecteerde mannen 'lijken' dus misschien meer op vrouwelijke directeuren dan de gemiddelde groep mannelijke directeuren.

Vrouwen blijken meer onderwijskundig gericht te zijn dan mannen, zowel wat betreft het hechten van belang aan de onderwijskundige taak, het zelf in portefeuille hebben ervan, de tijd die aan de taak wordt besteed, en de onderwijskundige activiteiten die zij verrichten; dit laatste sekseverschil blijkt uit de docentgegevens nog sterker dan uit de gegevens van de directeuren zelf. Vrouwen zijn dus meer dan mannen gericht op het primaire proces, op de onderwijsdoelen van de school; mannen delegeren de onderwijskundige portefeuille vaker. Enkele buitenlandse onderzoeken wezen al in

deze richting (zie par. 1). Ook voor het Nederlandse basisonderwijs werd enige evidentie gevonden voor het meer verrichten van onderwijskundige activiteiten door vrouwen: Van de Grift en Kurek-Vriesema (1990) vonden in het basisonderwijs op drie van hun zeven onderwijskundige schalen dat vrouwen meer activiteiten verrichten dan mannen, namelijk met betrekking tot het benadrukken van leerlingprestaties, het begeleiden van docenten bij zwak presterende leerlingen en het doen van klasseobservaties. De resultaten van het hier besproken onderzoek verklaren de grotere gerichtheid van vrouwen op het primaire proces ook geldig voor het voortgezet onderwijs in Nederland. Er worden zelfs op meer schalen sekseverschillen in onderwijskundige gerichtheid gevonden dan het geval was in het basisonderwijs. Dit is des te opmerkelijker, omdat in paragraaf 2 een aantal overwegingen werden gegeven die het weinig plausibel maken dat de buitenlandse resultaten ook geldig zouden zijn voor met name het voortgezet onderwijs in ons land. Er is bijvoorbeeld geen diplomering van schoolleiders vereist, overheidsbeleid legt de nadruk op de organisatorische kant van het schoolleiderschap en onderwijskundig leiderschap heeft in onze scholen een lage status. Daardoor zou weleens een nivellering van sekseverschillen in onderwijskundige gerichtheid kunnen optreden. In dit licht mag het een bijzonder resultaat genoemd worden dat ook in Nederlandse voortgezet onderwijs, ondanks dergelijke tegenwerkende contextfactoren, vrouwen meer onderwijskundige activiteiten tentoonspreiden dan mannen.

Mannen zijn meer dan vrouwen gericht op de secundaire taken in de school. Zij hebben vaker dan vrouwen de taak administratie en beheer zelf in portefeuille en besteden ook meer tijd hieraan evenals aan de relatie met de buitenwereld. Mannen hechten evenals de vrouwen wel meer belang aan onderwijskundige dan aan administratieve taken. De resultaten van Leithwood et al., (1990) en Gousha (1986) dat schoolleiders zich in de praktijk vooral bezighouden met administratieve taken, terwijl ze als gewenste praktijk aangeven bezig te willen zijn met onderwijskundige taken in de school (par. 1), wordt dus slechts bevestigd voor mannelijke directeuren. Bij vrouwen blijkt zo'n scheiding in reële en gewenste prak-

tijk nauwelijks op te treden. Wanneer onderzoeken slechts onder mannen worden verricht, blijkt dit te leiden tot opsporing van slechts de helft van de waarheid.

In het MBO wordt in tegenstelling tot in de beide andere onderzochte schoolsoorten geen enkel significant sekseverschil gevonden in het verrichten van onderwijskundige activiteiten, noch in het in portefeuille hebben van de onderwijskundige taak, noch in de tijdsbesteding eraan. Anderzijds blijken vrouwen in het MBO evenals in het AVO en het Dag/Avond onderwijs wel meer belang te hechten aan de uitoefening van onderwijskundige activiteiten dan mannen. We lijken hier naast sekseverschillen geconfronteerd te worden met schoolcultuurverschillen: blijkbaar is het voor vrouwen in het MBO in tegenstelling tot in de andere twee schoolsoorten niet mogelijk om hun wensen in praktijk te brengen, met andere woorden om hun potentiële gerichtheid op het primaire proces om te zetten in actief handelen. Een ander gegeven dat met verschillen in schoolcultuur zou kunnen samenhangen, is dat vrouwen en mannen volgens de docentgegevens het sterkst verschillen in onderwijskundige gerichtheid in het Dag/Avond onderwijs. Vrouwen zijn in deze schoolsoort op alle schalen en subschalen hogelijk significant onderwijskundiger gericht dan mannen.

Uit dit onderzoek komt naar voren dat *seks in combinatie met schoolcultuur* de verschillen in het leidinggeven van mannen en vrouwen bepaalt. Bij onderzoek naar sekseverschillen in leidinggeven moet dus onderscheid worden gemaakt in de mate waarin de cultuur waarin dat leidinggeven plaats vindt 'vrouwelijk' of 'mannelijk' is. De beoordeling hiervan kan plaats vinden op basis van de aantallen vrouwen en mannen die in de organisatie leiding geven, werken en leren, in combinatie met de historische ontwikkeling van de cultuur. Het Dag/Avond onderwijs, voortgekomen uit de zogenoemde 'moedermavo', kan zowel qua aantallen vrouwen als qua historie opgevat worden als een 'vrouwelijke' cultuur. We kunnen hier spreken over vrouwelijke directeuren in een oorspronkelijk door vrouwen ontwikkelde cultuur. Het MBO is volgens de gehanteerde criteria een schoolsoort met een 'mannelijke' cultuur: ze omvat relatief veel technische en

economische onderwijsrichtingen waaraan van ouds her door mannen leiding wordt gegeven. De MDGO-scholen met van oorsprong relatief veel vrouwelijke directeuren en leerlingen zijn naar aanleiding van fusies veelal ingebed in MBO-instellingen met oorspronkelijk mannelijke directeuren, leerlingen en culturen. In het AVO hebben altijd beduidend meer mannelijke directeuren de scepter gezwaaid, met uitzondering van de oude huishoudscholen, de latere LHNO's. Deze LHNO's zijn inmiddels bijna allemaal gefuseerd met scholen met een LTO- en/of LEAO-richting of zijn opgegaan in scholengemeenschappen. Het AVO is dus een van oorsprong 'mannelijke' cultuur, hoewel het zich ook onderscheidt van het MBO in die zin dat het om algemeen vormend onderwijs gaat en niet om beroepsvoorbereiding in voor een groot deel technische vakken. In de onderzoeksresultaten wordt de plaats van de drie schoolsoorten op de dimensie 'mannelijke - vrouwelijke cultuur' gereflecteerd: vrouwen zijn in het Dag/Avond onderwijs beduidend sterker gericht op het primaire proces dan mannen, nog sterker dan in het AVO, terwijl vrouwelijke directeuren in het MBO een even grote gerichtheid ten toon spreiden als hun mannelijke collega's.

Met dit onderzoek is het verband gelegd tussen de sekse van de schooldirecteur en de gerichtheid op het primaire proces. Wanneer deze gerichtheid van de schoolleider inderdaad een belangrijk kenmerk is voor de effectiviteit van scholen, zijn vrouwen effectiever dan mannen, althans wanneer zij hun leiderschap kunnen uitoefenen in een min of meer bij hun sekse passende cultuur. De 'mannelijke' cultuur op scholen lijkt een negatieve invloed te hebben op de effectiviteit van het schoolleiderschap van vrouwen. In toekomstig onderzoek naar effectieve scholen moet de sekse van de schoolleider als variabele worden betrokken teneinde het verband tussen deze variabele en schooleffectiviteit direct te onderzoeken in plaats van via het onderwijskundig leiderschap van de directeur. Het hier gepresenteerde onderzoek leidt voorlopig tot een waarschuwing. In het MBO is met de schaalvergrotingsoperaties de cultuur van de scholen 'mannelijker' geworden. Op dit moment vinden in het AVO en in het Dag/Avond onderwijs dezelfde soort ont-

wikkelingen plaats. Schaalvergroting en decentralisatie vergroten de noodzaak tot het leggen van nadruk op de organisatorische kant van het leiderschap; ze vergen eerder administratief dan onderwijskundig leiderschap. Dit zou ertoe kunnen leiden dat vrouwen in die schoolsoorten straks evenmin in staat zullen zijn als momenteel in het MBO het geval is om het soort leiderschap uit te oefenen dat vooralsnog de schooleffectiviteit ten goede lijkt te komen.

Literatuur

- Adkinson, J.A. (1981). Women in school administration: A review of the research. *Review of Educational Research*, 51, 311-343.
- Althaus, R.P., & Rubin, D. (1970). The computerized construction of a matched sample. *American Journal of Sociology*, 76, 325-346.
- Bonhof, G., Greeuw, P. de, Pollemans, M., & Vogel, E. (1986). *De laatste directrices*. Reeks: Beleid, Bestuur en Beheer van Onderwijs, Universiteit van Amsterdam.
- Bosman, R., & Louwes, W. (1982). *Eenoudergezin en schoolprestaties*. Rijksuniversiteit Groningen.
- Bossert, S.T., Dwyer, D.C., Rowan, R., & Lee, G.V. (1982). The instructional management role of the principal. *Educational Administration Quarterly*, 18(3), 34-64.
- Brandt, R. (1992). On rethinking leadership: a conversation with Tom Sergiovanni. *Educational Leadership*, 49(5), 46-49.
- Brubaker, D.L., & Simon, L.H. (1987). How do principals view themselves and others? *National Association of Secondary School Principals Bulletin*, 71(495), 72-82.
- Charters, W.W. Jr., & Jovick, T.D. (1981). The gender of principals and principal-teacher relations in elementary schools. In P.A. Schmuck, W.W. Charters Jr. & R.O. Carlson (Eds.), *Educational Policy and management: Sex differentials* (pp. 307-331). New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed., pp. 369-380). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dwyer, D.C., Lee, G.V., Barnett, B.G., Filby, N.N., Rowan, B., Alpert, B., & Kojimoto C. (1985). *Understanding the principal's contribution to instruction: Seven principals, seven stories* (Vols. 1-7 and Methodological Appendix). San Francisco, CA: Far West Laboratory.

- Dwyer, D.C., Barnett, B.G., & Lee, G.V. (1987). The school principal: scapegoat or the last great hope? In *Leadership, examining the elusive*. ASCD Yearbook (pp. 30-46).
- Eagly, A.H., & Johnson, B.T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin*, 108, 233-256.
- Eagly, A.H., Karau, S.J., & Johnson, B.T. (1992). Gender and leadership style among school principals: A meta-analysis. *Educational Administration Quarterly*, 28(1), 76-102.
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-24.
- Fishel, A., & Pottker, J. (1977). Performance of women principals: A review of behavioral and attitudinal studies. In J. Pottker & A. Fishel (Eds.), *Sex bias in the schools: The research evidence* (pp. 289-299). Cranbury, NJ: Associated University Presses, Inc..
- Fraser, J., & Fraser, R.S. (1979). Educational administration: A feminine profession. *Educational Administration Quarterly*, 15(2), 1-13.
- Gousha, R.P. (1986). *The Indiana school principalship: The role of the Indiana principal as defined by the principal*. Bloomington, Indiana: Indiana University School of Education.
- Grift, W. van de (1985a). Onderwijsleerklimaat en leerlingprestaties. *Pedagogische Studiën*, 62, 401-414.
- Grift, W. van de (1985b). *Het meten van onderwijskundig leiderschap*. Amsterdam: SCO.
- Grift, W. van de (1985c). *Vragenlijst onderwijskundig leiderschap, (OL)*. Amsterdam: SCO.
- Grift, W. van de (1987). Zelfpercepties van onderwijskundig leiderschap en gemiddelde leerlingprestaties. In J. Scheerens en W.G.R. Stoel (red.), *Effectiviteit van onderwijsorganisaties. Bijdragen aan de onderwijsresearch no.19* (pp. 83-98). Lisse: Swets & Zeitlinger B.V..
- Grift, W. van de, & Kurek-Vriesema, C. (1990). *Schoolleidingen in de basisschool*. Forum 1, Amsterdam/ Lisse: Swets & Zeitlinger B.V..
- Grobman, H., & Hines, V.A. (1956). What makes a good principal? *National Association of Secondary School Principals Bulletin*, 40, 5-16.
- Gross, N., & Trask, A.E. (1964). *Men and women as elementary school principals* (Final Report No.2). Cambridge, MA: Harvard University, Graduate School of Education.
- Gross, N., & Trask, A.E. (1976). *The sex factor and the management of schools*. New York: John Wiley.
- Hallinger, Ph. (1989a). *Developing instructional leadership teams in high schools*. Paper presented at the international congress for school effectiveness, Rotterdam, The Netherlands.
- Hallinger, Ph. (1989b). Developing instructional leadership teams in secondary schools: A framework. *National Association of Secondary School Principals Bulletin*, 84-92.
- Hallinger, Ph., Bickman, L., Davis, K. (1990). *Modeling the effects of principal leadership on student achievement*. Paper presented at the international congress for school effectiveness, Jerusalem, Israel.
- Hallinger, Ph., Murphy, J., Weil, M., Mesa R., & Mitman, A. (1983). School effectiveness, identifying the specific practices, behaviors for principals. *National Association of Secondary School Principals Bulletin*, 83-91.
- Hallinger, Ph., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, 86, 217-247.
- Hallinger, Ph., & Murphy, J. (1987). Assessing and developing principal instructional leadership. *Educational Leadership*, 45(1), 54-61.
- Hemphill, J.K., Griffiths, D.E., & Frederiksen, N. (1962). *Administrative performance and personality: A study of the principal in a simulated elementary school*. New York: Teachers College, Columbia University.
- Jong, A.M. de, & Doorne-Huiskes, J. van (1989). *Stijl van leidinggeven van vrouwen en mannen*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Krüger, M.L. (1994). *Sekseverschillen in schoolleiderschap*. Academisch proefschrift. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.
- Leithwood, K.A., Begley, P.T., & Cousins, J.B. (1990). The nature, causes and consequences of principals' practices: An agenda for future research. *Journal of Educational Administration*, 28(4), 5-31.
- Levinson, R.M. (1982). *Administrative task orientation of school principals*. Doctoral Dissertation. Santa Barbara: University of California.
- Mahieu, P. (z.j.). Management en leiding geven. In B. Creemers, J. Giesbers, C. van der Perre, C. van Vils-teren (red.), *Handboek voor Schoolorganisatie en Onderwijsmanagement* (pp. 4150/1-22). Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.
- Marx, E. (1986). Schoolorganisatorische aspecten van waardengerichtheid in het onderwijs bij uiteenlopende schoolculturen. In B. Creemers, J. Gies-

- bers, C. van der Perre, C. van Vilsteren (red.). *Handboek Schoolorganisatie en Onderwijsmanagement* (pp. 2110/1-27). Alphen aan den Rijn: Samsom Uitgeverij.
- Meester, A., & Leeuw, J. de (1983). *Intelligentie, sociaal milieu en de schoolloopbaan*. Rijksuniversiteit Leiden, FSW, Vakgroep Datatheorie.
- Meir, E.J. van, Jr. (1973). Leadership behavior of male and female elementary principals: A comparison by sex. *Marquette University Education Review*, 4, 8-11.
- Meskin, J.D. (1974). The performance of women school administrators: A review of the literature. *Administrators Notebook*, 23(1), 1-4.
- Meskin, J.D. (1979). Women as principals: Their performance as educational administrators. In D. Erickson & T.L. Reller (Eds.), *The principal in metropolitan schools*. Berkeley: McCutchan.
- Morsink, H.M. (1970). Leader behavior of men and women principals. *National Association of Secondary School Principals Bulletin*, 54, 80-87.
- Nunnally, J.C. (1978). *Psychometric theory* (2nd edition). New York/London: McGraw-Hill.
- Peschar, J.L. (1975). *Milieu - school - beroep*. Academisch proefschrift, Rijksuniversiteit Groningen.
- Ruijs, A. (1990). *Vrouwen en schoolmanagement*. Amsterdam/Lisse: Swets & Zeitlinger B.V..
- Ruijs, A., Mens, A., Baggen, C., & Jansen, B. (1986). *Schoolmanagement: een mannenzaak?* Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Ruijs, A., & Mientjes, A. (1989). *Werving en selectie van schoolmanagers*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Scheerens, J. (1989). *Wat maakt scholen effectief?* Samenvatting en analyse van onderzoeksresultaten. 's-Gravenhage: SVO.
- Shakeshaft, C. (1987). *Women in educational administration*. Beverly Hills: SAGE Publications, Inc..
- Sweeney, J. (1982). Research synthesis on effective school leadership. *Educational Leadership*, 39, 346-352.
- Vegt, A.L. van der, & Vos, W.A. de (1990). Onderwijskundig leiderschap. In B. Creemers, J. Giesbers, C. van der Perre, C. van Vilsteren (red.), *Handboek Schoolorganisatie en Onderwijsmanagement* (pp. B 1210/1-18). Alphen aan den Rijn: Samsom Uitgeverij.
- Yeakey, C.C., Johnston, G.S., & Adkinson, J.A. (1986). In pursuit of equity: A review of research on minorities and women in educational administration. *Educational Administration Quarterly*, 22(3), 110-149.

Auteur

M.L. Krüger werkt als universitair docent aan de Vakgroep Onderwijskunde, het Instituut voor de Lerarenopleiding en het Centrum voor Nascholing van de Universiteit van Amsterdam.

Adres: Universiteit van Amsterdam, Faculteit der Pedagogische en Onderwijskundige Wetenschappen, Wibautstraat 4, 1091 GM Amsterdam.

Abstract

Gender and school leadership in relation with the orientation to the primary processes

M.L. Krüger. *Pedagogische Studiën*, 1996, 73, 25-41.

It was not until the seventies that the gender variable came to be included in school management research. Moreover, the scarce research results are often inconsistent and sometimes even conflicting. This lack of research based knowledge has led to conducting a Ph.D research into gender differences in school principalship in Dutch secondary education. Ample attention has been paid to the orientation of school principals to the primary processes. This part of the research will be highlighted in this article. By means of a 'matching' method, pairs of compatible schools were generated, the 'only' difference being that the first school was managed by a female and the second by a male principal. 98 principals, 673 teachers and 858 students completed a questionnaire. The constructed instruments, mostly Likert-scales, were subjected to factor analyses and reliability measurements. The data were analyzed by means of pairwise multivariate analysis. The results demonstrate that the 'gender' variable has significant effects on leadership performance: women appear to be stronger oriented to the product of the organization, the primary processes, than their male counterparts. Apart from that, the results vary with different school cultures.

Het innovatief vermogen van scholen in het basisonderwijs: een tweede vooronderzoek

F. Geijssel, R. van den Berg, P. Slegers

Samenvatting

Scholen zullen steeds meer verantwoordelijkheid gaan dragen voor hun eigen toekomst. In dit kader is het zinvol om na te gaan welke potenties en mogelijkheden scholen hebben om innovaties en met name grootschalige, te realiseren. Hiertoe is een strategie voor onderzoek naar het innovatief vermogen van scholen opgezet.

In een eerdere publikatie in dit tijdschrift werd verslag gedaan van het eerste vooronderzoek dat in het voortgezet onderwijs heeft plaatsgevonden. Dit onderzoek heeft geleid tot een eerste conceptueel model van innovatief vermogen.

In deze bijdrage wordt ingegaan op het tweede vooronderzoek, in het basisonderwijs. De inhoud van de vernieuwing hierbij is 'Weer Samen Naar School' (WSNS). Dit onderzoek heeft tot doel meer inzicht te verwerven in het reeds ontwikkelde model van innovatief vermogen ten einde een vragenlijst te kunnen ontwikkelen voor grootschalig onderzoek. Bovendien is ervoor gekozen om een stap verder te gaan door de relatie tussen innovatief vermogen en betrokkenheid te onderzoeken. Voor dit onderzoek is een aantal 'hoog innovatieve' en 'laag innovatieve' scholen gezocht. Bij leerkrachten, interne begeleiders en directeurs ($N = 64$) van deze scholen zijn twee interviews en een vragenlijst afgenomen. De resultaten van de interviews en de vragenlijst hebben geleid tot meer inzicht in de betekenis van de componenten van innovatief vermogen.

den. Scholen komen steeds meer op eigen benen te staan, waardoor een beroep wordt gedaan op eigenschappen die betrekking hebben op de capaciteiten van de school als organisatie (Lagerweij, 1993). Scholen zullen zich steeds meer moeten ontwikkelen als strategische organisaties (Slegers, 1991; Van den Berg, 1992). Door een beroep te doen op individuele verantwoordelijkheid en creativiteit kunnen scholen een eigen richting kiezen en zelf invulling geven aan de vele eisen die de omgeving stelt. Om meer inzicht te krijgen in dergelijke processen is een lange termijn strategie opgezet voor onderzoek naar het innovatief vermogen, bestaande uit twee vooronderzoeken en een grootschalig promotie-onderzoek. Beide vooronderzoeken hebben tot doel het begrip innovatief vermogen te conceptualiseren. Eerder werd in dit tijdschrift het verslag gepubliceerd van het eerste vooronderzoek, dat plaats heeft gevonden in het voortgezet onderwijs (Van den Berg, Slegers, Bakx & Van der Eerden, 1994). In dit eerste onderzoek is een model van innovatief vermogen ontwikkeld, bestaande uit vier componenten: context, samenwerking van leerkrachten, transformatief leiderschap en de school als lerende organisatie.

In dit artikel wordt verslag gedaan van het onderzoek naar het innovatief vermogen in het basisonderwijs. Daarbij heeft de vernieuwing 'Weer Samen Naar School' (WSNS) centraal gestaan. Er zijn opnieuw interviews afgenomen met als doel nog meer inzicht te krijgen in de vier componenten. Bovendien is ervoor gekozen om een stap verder te gaan door de relatie tussen innovatief vermogen en betrokkenheid te onderzoeken. Betrokkenheid is onderzocht door gebruik te maken van het interview 'Gebruiksniveaus van Innovatie' en de vragenlijst 'Intensiteit van Betrokkenheid'. Sinds 1981 wordt er met deze instrumenten, ontleend aan het Betrokkenheidsmodel, onderzoek gedaan (Van den Berg & Vandenberghe, 1981, 1984,

1 Inleiding

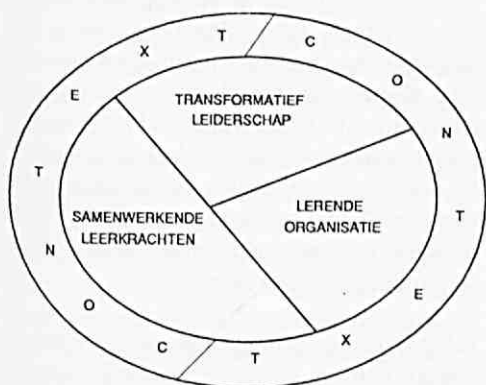
In de huidige maatschappij worden scholen voortdurend geconfronteerd met veranderingen in directe omgeving en vernieuwingen, waarvan vele door de overheid geïnitieerd wor-

1988, 1995; Kuijlaars, 1992; Ter Beek, 1992; Houtveen & Van de Grift, 1993; Van den Berg, 1993).

Dit artikel is als volgt opgebouwd. Allereerst wordt in paragraaf twee een korte schets gegeven van het innovatief vermogen van scholen. In paragraaf drie wordt het gebruikte onderzoeksdesign beschreven. Vervolgens geven we in de vierde paragraaf de resultaten weer. Dit doen we per component van het innovatief vermogen. We besluiten in paragraaf vijf met de conclusies. In de discussie zullen enkele resultaten worden geïnterpreteerd vanuit het perspectief van betrokkenheid.

2 Het innovatief vermogen van scholen

Bij de start van dit vooronderzoek is uitgegaan van het model en de definitie van het innovatief vermogen zoals door Bakx en Van der Eerden (1993) omschreven (zie ook Van den Berg et al., 1994). Het begrip 'innovatief vermogen' werd als volgt omschreven: *de competentie van scholen om zowel door de overheid geïnitieerde vernieuwingen, als door de school zelf op gang gebrachte vernieuwingen te implementeren en er bovendien zorg voor te dragen dat, indien nodig, beide soorten vernieuwingen in relatie tot elkaar staan.* Figuur 1 geeft het conceptueel model weer.



Figuur 1. Conceptueel model van innovatief vermogen

Op basis van literatuurstudie over het functioneren van innovatieve organisaties bleken de volgende vier componenten van belang voor

het innovatief vermogen (Bakx & Van der Eerden, 1993; Van den Berg et al., 1994): de context van de school, samenwerking van leerkrachten, transformatief leiderschap en het functioneren van de school als lerende organisatie. In het eerdere artikel in dit tijdschrift over het innovatief vermogen van scholen (Van den Berg et al., 1994) zijn deze componenten uitvoerig beschreven. We beperken ons hier tot de meest wezenlijke aspecten ervan.

Context

In dit kader wordt verwezen naar kenmerken van de externe omgeving van de school, waarbij markten, politiek klimaat en economische omstandigheden een rol spelen (Mintzberg, 1991). De context van de school is momenteel sterk aan het veranderen, waardoor scholen gedwongen worden te innoveren. Daarbij worden de mogelijkheden van scholen begrensd door het vigerende beleid van het Ministerie van Onderwijs, Cultuur en Wetenschappen. Dit beleid is de laatste jaren gericht op een vergroting van de autonomie, waardoor de vrijheid van handelen van scholen toeneemt. Wil een school haar innovatief vermogen vergroten, dan zal ze optimaal gebruik moeten maken van deze geboden vrijheid. Van belang hierbij is de manier waarop scholen deze context percipiëren. Hoe groot schatten scholen de beschikbare beleidsruimte in en wat doen ze ermee (Rolff, 1993; Slegers & Wesselingh, 1993; Vandenbergh, 1993)?

Samenwerking van leerkrachten

Een goed functionerend team van leerkrachten is het bouw materiaal van de school (Van den Berg, 1990, 1992). Het team bepaalt het functioneren van de organisatie en ook het innovatief vermogen ervan. De schoolleiding speelt daarbij een voorwaardenscheppende rol (Lagerweij, 1993; Lagerweij & Haak, 1994). Innovatie vereist dat het team niet als een groep individuele docenten functioneert, maar dat de krachten van docenten gebundeld worden (Mintzberg 1991). Het bevorderen van samenwerking tussen docenten, zowel binnen als tussen de vaksecties, vergroot de onderlinge steun en collectieve verantwoordelijkheid. Binnen deze collegialiteit moet er ruimte zijn voor individuele vrijheid en creativiteit. Het vinden van een evenwicht tussen collegialiteit aan de

ene kant en individuele vrijheid en creativiteit aan de andere kant is essentieel voor de ontwikkeling van het innovatief vermogen (Van den Berg et al., 1994).

Transformatief leiderschap

Een vorm van leiderschap die aansluit bij het ontwikkelen van een hoog innovatief vermogen van de school wordt transformatief leiderschap genoemd. In het eerste vooronderzoek is nagegaan door welke aspecten transformatief leiderschap wordt gekenmerkt. Deze zijn: visieontwikkeling, het stimuleren van samenwerking en het respecteren van de zelfontwikkeling van leerkrachten (zie ook: Leithwood, 1994). Vooral bij grootschalige vernieuwingen vervult de transformatieve leider een belangrijke rol. Transformatieve schoolleiders streven drie fundamentele doelen na (Leithwood, Tomlinson & Genge, 1995):

- 1 Het stimuleren en ontwikkelen van een gemeenschappelijk klimaat;
- 2 Een bijdrage leveren aan de professionele ontwikkeling van docenten;
- 3 Uitbreiding van het probleemoplossend vermogen van de school.

Transformatief schoolleiderschap vindt plaats binnen een kader waarin de potenties van de betrokkenen in een school worden gemaximaliseerd. De eigen verantwoordelijkheid van individuen speelt hierbij een belangrijke rol (Van den Berg, 1990). Hierbij is sprake van een complementariteit tussen enerzijds de eigen ontwikkelingsbehoefte van docenten en anderzijds het scheppen van mogelijkheden door de organisatie en de leider voor het realiseren van vernieuwingen. Waar grootschalige vernieuwingen leiden tot ontwrichting van bestaande verwachtingspatronen bij docenten (Salisbury & Conner, 1994), is transformatief leiderschap nodig met als streven docenten naar hogere niveaus van betrokkenheid en motivatie te brengen.

De school als lerende organisatie

Een lerende organisatie is een "organisatie die als bewust beleid voert haar lerend vermogen op alle niveaus en op continue basis te vergroten, ter optimalisering van haar effectiviteit" (Bomers, 1989, p. 2). Volgens Louis (1994) past de theorie rondom de lerende organisatie in het denken over vernieuwingen als non-

lineaire en niet-rationele processen. Binnen het concept van de lerende organisatie wordt de nadruk gelegd op cognitieve en gedragsmatige transformatie, zowel bij individuen als bij groepen (Louis, 1994, p. 8). Marx en De Vries (1995, p. 61) zien de lerende organisatie als een complexe organisatie, waarin diverse strijdige opties, zoals individuele autonomie en wederzijdse afhankelijkheid, met elkaar in evenwicht worden gebracht. Om het leren van organisaties als geheel mogelijk te maken zijn samenwerking en interactie van belang.

Dit organisatie-leren is meer dan de optelsom van het leren van individuen. Senge (1993) doelt op hetzelfde met 'teamleren' als een van de vijf disciplines die hij noodzakelijk acht voor het functioneren van een lerende organisatie. In relatie tot dit teamleren noemt hij de noodzaak van een gemeenschappelijke visie om richting te geven aan het leerproces in de samenwerking van mensen. Coördinatie binnen de lerende organisatie vindt dan voornamelijk plaats op basis van onderlinge afstemming tussen de verschillende geledingen. De basisstructuur van een organisatie waarin dit coördinatie-mechanisme centraal staat, is een organische structuur. In veel literatuur over innovatieve organisaties (Mintzberg, 1991; Cobbenhage, 1990; Sundbo, 1992; e.a.) wordt een organische structuur gezien als een adequate organisatievorm om in een complexe en dynamische omgeving te kunnen overleven. Een ander kenmerk van een organische structuur is strategievorming als beleidsbepalend instrument (Morgan, 1986; Mintzberg, 1991). Beleidsbepaling vindt niet plaats door strategische planning op één plaats aan de top (het management). Beleid wordt veel meer op impliciete wijze geformuleerd doordat zich allerlei specifieke activiteiten in de organisatie afspelen. Ontwikkeling van beleid krijgt hierdoor een incrementeel karakter: beleid groeit stap voor stap, vanuit de basis.

In het kader van de implementatie van vernieuwingen worden, naast het innovatief vermogen, ook begrippen als beleidsvoerend vermogen (Sleegers, 1991) en veranderingscapaciteiten (Van Gennip, 1991) van belang geacht. In het eerdere artikel in dit tijdschrift over het innovatief vermogen van scholen (Van den Berg et al., 1994) beschreven we dat deze

concepten vooral organisatorische aspecten van de school betreffen, terwijl het innovatief vermogen ook culturele aspecten omvat. Lagerweij en Haak (1994) gebruiken het begrip schoolontwikkeling als perspectief voor onderwijsverandering. Zij onderscheiden daarbij een structurele en dynamische invalshoek (p. 98). Het begrip innovatief vermogen sluit aan bij de laatstgenoemde invalshoek, en wel bij de vraag hoe de verschillende elementen van het proces veranderbaar zijn. Het innovatief vermogen richt zich op de mensen en hun ontwikkeling in de school die zich als een lerende organisatie profileert.

Betrokkenheid

Bij bovenstaande beschrijving van de componenten van het innovatief vermogen zijn op meerdere plaatsen elementen van het 'betrokkenheidsdenken' te herkennen. Dit blijkt met name uit het belang van de individuele verantwoordelijkheid en het persoonlijk perspectief van docenten voor het innovatief vermogen van scholen. Met betrokkenheid bedoelen Van den Berg en Vandenberghe (1995) de *vragen* met een min of meer *emotionele* ondertoon, die *onzekerheid*, eventueel *weerstand* tegen verandering signaleren en die op één of andere wijze door de leerkracht(en) zullen moeten worden opgelost. Van den Berg en Vandenberghe (1981, 1984, 1988, 1995) onderscheiden drie hoofdvarianten van betrokkenheid: zelf-, taak- en anderbetrokkenheid. Betrokkenheid is een natuurlijk fenomeen bij vernieuwingsprocessen en kan beschouwd worden als iets dat zich bij veranderingen in de dagelijkse praktijk manifesteert (zie ook Fullan & Stiegelbauer, 1991).

Vooraf in de laatste decennia rust een permanente druk tot voortdurende verandering. Deze permanente druk komt voort uit de spanning tussen de nieuwe opdrachten die de school uit de samenleving krijgt, en de dikwijls behoudende cultuur op scholen. Daar komt nog bij dat vernieuwingen vaak grootschalig zijn en door de landelijke overheid aan de scholen worden opgelegd. Iedere leerkracht laat zich echter leiden door zijn of haar eigen opvattingen over onderwijs geven, zijn of haar subjectieve onderwijsstheorie (Van den Berg & Vandenberghe, 1995; Huberman, 1993). Betrokkenheid kan men op deze manier ook be-

schouwen als een proces waarbij leerkrachten hun subjectieve onderwijsstheorie confronteren met uiteenlopende verwachtingen vanuit een wisselende omgeving. Ook Senge (1993) ziet de individuele motivatie van organisatieleden ('persoonlijk meesterschap') als bepalend voor een goed functioneren van de organisatie. Het gaat om een wisselwerking tussen individu en organisatie als een proces van schoolontwikkeling, dat altijd doorgaat. Zo beschouwd lijken betrokkenheid en het innovatief vermogen van scholen aan elkaar gerelateerd. Deze veronderstelling is in de hier gerapporteerde studie onderzocht.

3 Het onderzoeksdesign

3.1 Onderzoeksvragen

De volgende vragen stonden centraal in het onderzoek:

- Welke concrete verschillen bestaan er per component van het innovatief vermogen tussen enerzijds hoog innovatieve en anderzijds laag innovatieve basisscholen?
- Is er een relatie tussen innovatief vermogen van basisscholen en de mate van betrokkenheid bij een vernieuwing?

Uit het eerste vooronderzoek is gebleken dat een groep hoog innovatieve scholen in het voortgezet onderwijs verschilde van een groep laag innovatieve scholen voor wat betreft de genoemde componenten. In dit tweede vooronderzoek willen we dieper op deze verschillen ingaan. Ten aanzien van de tweede onderzoeksvraag wordt verwacht dat op de hoog innovatieve scholen een grotere betrokkenheid wordt gevonden dan op de laag innovatieve scholen.

3.2 Onderzoekopzet

Voor dit onderzoek zijn scholen gezocht met een hoog, dan wel laag innovatief vermogen. Gerards (1993) ontleent aan een aantal eerdere onderzoeken naar vernieuwingscholen de volgende drie criteria die gebruikt kunnen worden om scholen op basis van het innovatief vermogen te selecteren:

- het aantal uitgevoerde vernieuwingsprojecten in het verleden;
- het aantal vernieuwingsgebieden in uitvoering;

c het oordeel van de betreffende begeleider uit de regio.

De eerste twee criteria hebben als assumptie dat scholen met een laag innovatief vermogen in het verleden minder vernieuwingsprojecten hebben uitgevoerd en momenteel op minder vernieuwingsgebieden bezig zijn dan scholen met een hoog innovatief vermogen. Deze criteria kunnen vanzelfsprekend en valide zijn, maar roepen ook vragen op. Zo zou men zich bijvoorbeeld een school kunnen voorstellen die nog maar vijf jaar geleden opgericht is en toch zeer innovatief te werk gaat. Vandaar dat gekozen is om voor de selectie van basisscholen aan het derde criterium prioriteit te geven.

Om hoog en laag innovatieve scholen te vinden hebben gesprekken plaatsgevonden met drie schoolbegeleiders uit drie regio's, die zich in hun werk concentreren op WSNS. Aan elke schoolbegeleider is gevraagd twee hoog innovatieve en twee laag innovatieve scholen te zoeken voor deelname aan het onderzoek, uitgaande van de vorderingen binnen het project WSNS. De vier door ons geschetste componenten van innovatief vermogen hebben bij de uiteindelijke selectie van de scholen geen rol van betekenis gespeeld. Wel is met de begeleiders gesproken over de mogelijke betekenis van innovaties voor scholen; dit in zeer algemene zin en in relatie tot de drie boven geschetste criteria. Het vinden van scholen, die naar de mening van de begeleider weinig investeerden in WSNS, leverde de meeste problemen op. Uiteindelijk hebben vier van deze scholen meegedaan. Scholen met een grote investering in WSNS (naar het oordeel van de begeleider) zijn eerder bereid tot medewerking. Het was dan ook geen probleem om zes hoog innovatieve scholen te vinden. In totaal hebben dus tien scholen meegewerkt aan dit onderzoek.

Aan de schoolleider van elke school is een korte vragenlijst ter invulling toegestuurd. Via deze vragenlijst is onder andere informatie verkregen over uitgevoerde vernieuwingsprojecten in het verleden en huidige vernieuwingsgebieden (criterium a en b). Het bleek dat in het verleden en op het moment van onderzoek op de zes hoog innovatieve scholen sprake was van relatief meer vernieuwingsprojecten dan op de vier laag innovatieve scholen. Dit ondersteunt de door Gerards aangegeven criteria.

Op de tien deelnemende scholen hebben de

directeur, de interne begeleider en leerkrachten meegedaan aan het onderzoek ($N = 64$). Bij deze 64 respondenten zijn twee interviews en een vragenlijst afgenomen. Het innovatief vermogen is met behulp van een interview onderzocht, dat een aangepaste versie is van het interview dat in het voortgezet onderwijs is afgenomen (Bakx & Van der Eerden, 1993). De vragen hadden betrekking op de vier componenten van het ontwikkelde model.

De betrokkenheid bij de vernieuwing is onderzocht met het interview 'Gebruiksniveau van Innovatie' (Van den Berg & Vandenbergh, 1981) en de Rasch-vragenlijst 'Intensiteit van Betrokkenheid' (Houtveen & Van de Grift, 1993). Met het interview kan worden achterhaald in welke fase van een vernieuwing een leerkracht zich bevindt. Met de vragenlijst kan worden vastgesteld hoe groot de intensiteit van de betrokkenheid bij een vernieuwing is aan de hand van Rasch-scores die op basis van de ingevulde vragenlijst kunnen worden bepaald.

Bij de interviews en de vragenlijst heeft de vernieuwing 'Weer Samen Naar School' (WSNS) centraal gestaan. Om de verschillende scholen en respondenten in dit onderzoek met elkaar te kunnen vergelijken is het van belang te waarborgen dat alle respondenten én de onderzoeker zo veel mogelijk dezelfde inhoud aan WSNS toekennen. Hierbij is rekening gehouden met het verschil in perspectief van de directeuren, de interne begeleiders en de leerkrachten. Bij de leerkrachten betrof de omschrijving de concrete uitvoering van WSNS in de dagelijkse praktijk van het lesgeven. Bij de directeuren had de omschrijving betrekking op het ontwikkelen van beleid ten aanzien van WSNS voor hun school en in het samenwerkingsverband. Bij de interne begeleiders betrof de omschrijving hun taken met betrekking tot het operationaliseren van WSNS-beleid in hun school (Geijssels, 1994). Voor de interpretatie van de onderzoeksresultaten is het van belang rekening te houden met dit verschil in perspectieven.

3.3 Analyse

Alle interviews 'Innovatief vermogen van scholen in het basisonderwijs' zijn op band opgenomen en volledig uitgewerkt door derden. Deze interviews vormen het ruwe materiaal

van dit onderzoek. Vervolgens hebben de onderzoekers de interviews opgedeeld in kleine teksteenheden. Elke teksteenheid bevat een of meerdere uitspraken van de geïnterviewde met betrekking tot een bepaalde gebeurtenis of geeft een gedachte weer van de geïnterviewde. Deze teksteenheden zijn vervolgens met elkaar vergeleken en ondergebracht bij een beperkter aantal categorieën. In totaal zijn op deze wijze 1671 teksteenheden toegekend aan 65 categorieën. Veel van deze categorieën vormen tweetallen die als elkaars tegengestelde zijn te beschouwen. Van elk tweetal zou idealiter gesteld kunnen worden dat de ene categorie alleen bij hoog innovatieve scholen wordt aangetroffen, terwijl de andere categorie alleen op scholen met een laag innovatief vermogen wordt aangetroffen. Er is ook een aantal categorieën gevonden zonder tegengestelde categorie (zie hiervoor de Tabellen in par. 4).

Voor het categoriseren van de teksteenheden is gebruik gemaakt van het programma Teksttabel, een gecomputeriseerde methode voor het analyseren van tekstuele data (Welten & Janssen, 1993). De categorieën, die met behulp van het programma Teksttabel zijn gevonden, werden op hun beurt ondergebracht bij één van de vier componenten uit het model van innovatief vermogen. Dit is grotendeels gebeurd op basis van beschikbare literatuur over het innovatief vermogen (Bakx & Van der Eerden, 1993).

3.4 Betrouwbaarheid van de analyse

Bij het doen van kwalitatief onderzoek is de intersubjectiviteit tussen de onderzoekers een belangrijke methode om de betrouwbaarheid te vergroten. Met intersubjectiviteit wordt in dit verband bedoeld: consensus tussen de onderzoekers. De interviews naar het innovatief vermogen in dit onderzoek zijn door twee onderzoekers geanalyseerd. Smaling (1992) beschrijft drie traditionele vormen en twee meer alternatieve vormen van intersubjectiviteit. Aan elk van deze vormen is op een of andere wijze voldaan, zoals nu kort beschreven zal worden. Met *consensual intersubjectivity* wordt bedoeld de consensus tussen onderzoekers over het interpreteren van de data. Tijdens de analyse is in dit onderzoek voortdurend gestreefd naar consensus. De analyse van de teksteenheden van het interview naar het inno-

vatief vermogen is uitgevoerd door twee onderzoekers. Eén van hen was vanaf het begin bij het onderzoek betrokken geweest, de ander niet. De onderzoekers hebben de analyse samen uitgevoerd, waarbij beiden evenveel inbreng hadden. Ten aanzien van elke teksteenheid is consensus bereikt. De indeling van de categorieën volgens de vier componenten van het model is een iteratief proces geweest, waarbij tevens beide begeleiders van het onderzoek betrokken waren. Ook hierover is consensus bereikt.

Intersubjectivity by regimentation betreft het standaardiseren van dataverzamelingstechnieken om de individuele verschillen tussen de onderzoekers te minimaliseren. Hiertoe zijn de eerste zes interviews door de twee interviewers in elkaars bijzijn afgenomen. Telkens stelde één interviewer de vragen terwijl de ander kritisch luisterde naar de interviewer. Na elk interview heeft vervolgens een evaluatie plaatsgevonden. Op deze wijze hebben de beide interviewers zich ervan kunnen overtuigen dat zij op vergelijkbare wijze interviewden. Dit werd mede mogelijk door de afspraak in alle interviews dezelfde vragen te stellen.

De *intersubjectivity by explicitness* doelt op een zo eenduidig mogelijke beschrijving van de gebruikte materialen, methoden en onderzoeksontwerpen, en tevens op de argumenten voor keuzen die daarin gemaakt zijn. Hieraan is voldaan door in het verslag expliciet te beschrijven en te verantwoorden welke stappen er in het onderzoek genomen zijn (Geijsel, 1994).

Een minder gangbare vorm is de *argumentative intersubjectivity* die voortkomt uit de opvatting dat de wetenschap alleen nog voortgang kan boeken door samenwerking en discussie tussen onderzoekers. Smaling (1992) acht hierbij gelijkwaardigheid tussen de onderzoekers van groot belang en noemt onder andere de methodologische discussie. In dit onderzoek is van beide sprake geweest.

Dialogical intersubjectivity is tevens een minder gebruikelijke vorm. Hier moet niet alleen overeenstemming zijn tussen de onderzoekers onderling maar ook met de onderzochten zelf. Dit kan worden bereikt door het terugkoppelen van de resultaten aan de respondenten en het verwerken van eventuele reacties daarop. Voor elk van de tien scholen is een terugkoppelingsrapportage geschreven. Alle

scholen zijn in gelegenheid gesteld om toelichting op de rapportage te verkrijgen. Hiervan heeft één school gebruik gemaakt.

4 Resultaten

In de volgende subparagrafen worden de resultaten voor elke component van het innovatief vermogen afzonderlijk gerapporteerd. Daarbij wordt onderscheid gemaakt tussen de twee groepen scholen. De resultaten worden weer gegeven in Tabel 1 tot en met 4. Links bovenaan staat steeds de betreffende component genoemd. De categorieën van elke component zijn daaronder geplaatst (verticaal). Voor een groot gedeelte van de categorieën geldt dat deze op bepaalde wijze met elkaar te maken hebben of elkaars tegengestelde zijn. Dergelijke paren van categorieën worden in de tabellen telkens direct onder elkaar geplaatst en afgeschermd door horizontale lijnen.

In elke tabel wordt (horizontaal) de groep hoog innovatieve scholen ($N = 6$) vergeleken met de groep laag innovatieve scholen ($N = 4$). In de cellen van de tabel staat het aantal toegekende teksteenheden per categorie voor de groep hoog en de groep laag innovatieve scholen. Er is voor gekozen om de aantallen eenheden van de laag innovatieve scholen in deze tabel te vermenigvuldigen met een bepaalde factor, daar het aantal respondenten van de hoog en de laag innovatieve scholen niet aan elkaar gelijk zijn. De factor is de verhouding tussen het aantal respondenten op de hoog en de laag innovatieve scholen. Deze is: 45/19 (ongeveer 2.37). Door het absolute aantal toegekende teksteenheden van de groep laag innovatieve scholen met deze factor te vermenigvuldigen is het mogelijk zicht te krijgen op de verschillen tussen de twee groepen van scholen. Doel van deze kwalitatieve analyse is namelijk extremen met elkaar te vergelijken om tendensen op te sporen.

Bij het vergelijken van de groepen van scholen is het van belang te weten in hoeverre het aantal eenheden verspreid is over de scholen, zodat voorkomen wordt dat er uitspraken gedaan worden over een groep van scholen waar het feitelijk één school betreft. In dit onderzoek is daarom bekeken of er grote verschillen zijn tussen de scholen door de absolute aantallen

per school te vergelijken. Daar waar er verschillen tussen de scholen zijn, zal dit worden beschreven.

4.1 Context

In Tabel 1 is te zien hoeveel eenheden er per groep van scholen zijn toegekend aan de categorieën die horen bij de component context. Voor de exacte omschrijving van de categorieën wordt verwezen naar Geijssels (1994).

Tabel 1

Resultaten m.b.t. de component context per groep van scholen

CONTEXT	hoog innovatief	laag innovatief
voldoende beleidsruimte	14	0
onvoldoende beleidsruimte	2	21
samenwerkingsverband actief	8	7
samenwerkingsverband passief	3	14
WSNS ingevoerd	14	0
WSNS uitdaging	23	9
WSNS concreet	20	9
WSNS minimaal	9	45

Uit deze tabel blijkt dat de zes hoog innovatieve scholen ruimte zien om eigen beleid te voeren binnen de regelgeving van het ministerie. Op de vier laag innovatieve scholen blijkt daarentegen dat men zich zeer belemmerd voelt door het ministerie en geen mogelijkheden zoekt om een eigen beleid te voeren.

Ten aanzien van het samenwerkingsverband tussen basisscholen en het speciaal onderwijs in het kader van WSNS is het verschil minder scherp. Wellicht hebben leerkrachten een minder duidelijk beeld van wat er in het samenwerkingsverband zoal gebeurt. Bij de categorieën met betrekking tot WSNS is er wel weer verschil tussen beide groepen scholen. Op de hoog innovatieve scholen zijn in verhouding tot de laag innovatieve scholen meer uitspraken gedaan waaruit blijkt dat men van mening is dat er op de school al volgens de eisen van WSNS wordt gewerkt. Men vindt tevens dat WSNS uitdagingen en mogelijkheden biedt en men is in staat concreet aan te geven welke nieuwe taken WSNS met zich meebrengt. Op de laag innovatieve scholen zijn er, juist in tegenstelling tot de hoog innovatieve scholen, uitspraken gedaan waaruit blijkt dat WSNS met zo min mo-

gelijk consequenties voor de dagelijkse praktijk wordt ingevoerd.

Samengevat zijn er binnen de component context verschillen tussen de twee groepen van scholen gevonden in de mate waarin een school beleidsruimte percipiëert en de manier waarop men bezig is met en denkt over WSNS.

4.2 Samenwerking van leerkrachten

Tabel 2

Resultaten m.b.t. de component samenwerking van leerkrachten per groep van scholen

SAMENWERKING VAN LEERKRACHTEN	hoog innovatief	laag innovatief
samenwerking algemeen voldoende	21	19
samenwerking algemeen onvoldoende	3	14
samenwerking actief	21	2
samenwerking passief	1	55
harmonie	41	2
disharmonie	0	36
behoefte aan klasoverstijgende samenwerking	38	0
eilandencultuur	1	28
lage drempel	66	19
hoge drempel	7	50
tussen bouwen voldoende	7	0
binnen bouwen voldoende	13	2
tussen bouwen onvoldoende	4	7

Tabel 2 laat zien dat op hoog innovatieve scholen de mate van samenwerking in het algemeen voldoende wordt gevonden, terwijl op de laag innovatieve scholen de meningen meer verdeeld zijn. Het verschil is niet zo groot. Ook blijkt uit de gegevens dat er binnen de hoog innovatieve scholen meer sprake is van actieve, klasoverstijgende samenwerking, terwijl op de laag innovatieve scholen in verhouding meer sprake is van passieve samenwerking waarbij men de neiging heeft zich vast te houden aan bestaande structuren. De hoog innovatieve scholen kennen een harmonieuze sfeer. Op de laag innovatieve scholen blijkt daarentegen uit relatief veel uitspraken dat er veel minder sprake is van een onderlinge vertrouwensbasis en dat de sfeer niet altijd op elk front even open en vriendelijk is (disharmonie). Relatief veel respondenten van de laag innovatieve scholen geven dan ook aan dat er sprake is van een eilandencultuur in de school. Op hoog innova-

tieve scholen zijn zeer veel opmerkingen gemaakt, meerdere per respondent, waaruit blijkt dat de drempel om met elkaar over het onderwijs en onderwijsproblemen te praten laag is. Bij de laag innovatieve scholen daarentegen is er overwegend meer sprake van een hoge dan van een lage drempel. Daarnaast wordt op de hoog innovatieve scholen veel meer vorm gegeven aan het zogenaamde bouwoverleg. Met een 'bouw' wordt een afdeling in een basisschool bedoeld, waarbij vaak onderscheid gemaakt wordt tussen onderbouw (kleutergroepen), middenbouw (tot groep zes of zeven) en bovenbouw (de hoogste groepen).

Uit deze resultaten blijkt dat er op hoog innovatieve scholen meer wordt samengewerkt dan op de laag innovatieve scholen. Verder komt naar voren dat de groepen van scholen verschillen op aspecten als actieve of passieve samenwerking, de mate van harmonie, de behoefte aan klasoverstijgende samenwerking, de mate waarin er sprake is van een eilandencultuur en het al dan niet aanwezig zijn van een drempel om met elkaar te praten over het werk.

4.3 Transformatief leiderschap

Tabel 3

Resultaten m.b.t. de component transformatief leiderschap per groep van scholen

TRANSFORMATIEF LEIDERSCHAP	hoog innovatief	laag innovatief
visie	38	17
onvoldoende visie	1	43
charismatische gedrevenheid	22	0
onvoldoende charismatische gedrevenheid	0	38
innovatief	62	7
klassikaal onderwijs	0	26
individuele feedback voldoende	27	5
individuele feedback onvoldoende	4	12
adequaate delegeren	48	5
niet adequaat delegeren	12	38
gezamenlijke doelen	41	2
cultuur van samenwerking	41	2
gezamenlijke verantwoordelijkheid	7	2
geen gezamenlijke doelen	0	47

Uit Tabel 3 blijkt het volgende. Op hoog innovatieve scholen kenmerken de schoolleiders zich door hun visie op het onderwijs en de

school. Zij hebben een bepaald charisma, waardoor collega's geïnspireerd worden, en zij zijn in staat doortastend op te treden. Hiermee werkt de schoolleider voortdurend aan de realisering van haar of zijn streefbeeld. De schoolleider initieert bewust vernieuwingen in de school en maakt gebruik van overtuigingskracht om het team te motiveren. Daarbij heeft de schoolleider respect voor ieders ideeën met betrekking tot de vernieuwing en is er begrip voor persoonlijke behoeften en gevoelens. De schoolleider is in staat verantwoordelijkheden uit handen te geven zonder deze uit het oog te verliezen. Er wordt veel aandacht besteed aan het creëren van een gezamenlijk draagvlak. Zo is samenwerking vanzelfsprekend en dit wordt gestimuleerd door onder andere de initiatie van werkwijzen, zoals het bouwoverleg, waarin de gestelde doelen tot uitdrukking komen. Daarbij wordt het team door de schoolleider aangesproken op de gezamenlijke verantwoordelijkheid die het draagt naar de leerlingen en hun ouders.

Wanneer we kijken naar de laag innovatieve scholen, ontstaat een heel ander beeld van de schoolleider. Veel respondenten op deze scholen zijn van mening dat hun schoolleider onvoldoende onderwijsvisie heeft. Overigens werden vragen naar de visie van de schoolleider veelal beantwoord met 'geen idee' of 'ik begrijp de vraag niet'. De schoolleider is in veel mindere mate een initiator; zij of hij motiveert en inspireert het team niet. Leerkrachten laten blijken behoefte te hebben aan meer doortastendheid zoals het doorhakken van knopen bij langdurige discussiepunten (overigens zonder dat dit ten koste gaat van hun eigen plannen). De schoolleider van de laag innovatieve school staat achter het vastgestelde, dikwijls meer behoudende karakter van het onderwijs in de school en kiest er bewust voor bepaalde vernieuwingen zo veel mogelijk af te remmen. Op persoonlijke behoeften en gevoelens wordt veelal niet ingegaan. Delegeren wordt nagelaten of juist gezien als een manier om werk aan anderen over te laten. De schoolleider zet zich weinig in om een draagvlak te creëren voor veranderingen en vernieuwingen. In veel mindere mate dan op de hoog innovatieve scholen is er iets 'eigens' aan de school, waar men gemeenschappelijk voor staat. Een van de laag innovatieve scholen vormt op dit beeld enigszins een uitzondering. Een jaar geleden is daar een

nieuwe schoolleider benoemd die zich onder andere ten doel stelt de samenwerking in de school te verbeteren.

Samengevat zien we verschillen tussen het leiderschap in de twee groepen van scholen wat betreft de ontwikkeling van een visie, het nastreven van gezamenlijke doelen, het inspireren van teamleden en begrip hebben voor persoonlijke gevoelens en behoeften.

4.4 De school als lerende organisatie

Tabel 4

Resultaten m.b.t. de component lerende organisatie per groep van scholen

LERENDE ORGANISATIE	hoog innovatief	laag innovatief
invloed gestimuleerd	70	14
invloed onvoldoende gestimuleerd	6	38
taakbeschrijving duidelijk	19	12
taakbeschrijving onduidelijk	5	14
lopende zaken bekend	40	26
lopende zaken vaak gemist	5	14
beschikbaarheid van informatie voldoende	32	7
beschikbaarheid van informatie onvoldoende	9	36
scholingsbeleid voldoende	20	2
scholingsbeleid onvoldoende	3	12
scholingsbehoefte groot	31	7
minimale scholingsbehoefte	14	43
besluiten in teamvergadering	57	12
weinig inspraak bij besluiten	2	24
geen eensgezindheid bij besluiten	0	21
middenmanagement	33	7
persoonlijke groei-behoefte	32	14

Uit de aantallen in Tabel 4 blijkt dat op hoog innovatieve scholen in verhouding meer invloed en initiatief vanuit het team wordt gestimuleerd dan op de laag innovatieve scholen. Dit duidt op het bestaan van strategievorming. Taakbeschrijvingen van bijvoorbeeld de interne begeleider zijn relatief minder vaak onduidelijk en lopende zaken worden relatief minder vaak gemist. Tussen de twee groepen van scholen is er in verhouding een duidelijk verschil ten aanzien van de beschikbaarheid van informatie. Zo is op de hoog innovatieve school de orthotheek veelal beter ingericht en is

men op de hoogte van wat er aan materiaal in de school aanwezig is. Uitspraken van respondenten geven blijk van scholingsbeleid, terwijl op de laag innovatieve scholen relatief veel minder wordt nagedacht over bijvoorbeeld welke cursussen in de toekomst voor de school van belang zouden kunnen zijn. Ook wordt in verhouding met de groep hoog innovatieve scholen veel vaker uiting gegeven aan een minimale scholingsbehoefte. Men vindt het minder belangrijk om op de hoogte te blijven van nieuwe ontwikkelingen.

Ook bij het nemen van besluiten zijn er verschillen. Op de hoog innovatieve scholen gebeurt dit in de teamvergadering, waarbij gestreefd wordt naar consensus; dit geldt met name voor beleidsbeslissingen. Op de laag innovatieve scholen worden besluiten ook wel in teamverband genomen, maar inspraak wordt relatief veel minder gestimuleerd. Uit veel van de uitspraken blijkt dat beslissingen wel in consensus gebeuren, tenzij wanneer de meningen sterk verdeeld zijn. Verder zijn er op de hoog innovatieve scholen meer teksten toegekend aan de categorieën middenmanagement en persoonlijke groei-behoefte. Blijkbaar is er op deze scholen relatief meer sprake van een middenkader binnen de organisatie dan op de laag innovatieve scholen. Uitspraken in de categorie 'persoonlijk groei-behoefte' geven aan dat de respondent er naar streeft zich op allerlei vlakken te blijven ontwikkelen en kritisch te blijven op zijn of haar functioneren.

Verschillen tussen de twee groepen van scholen voor wat betreft de lerende organisatie, betreffen dus vooral de invloed van teamleden, de beschikbaarheid van informatie, het scholingsbeleid en de manier waarop belangrijke besluiten worden genomen.

4.5 Betrokkenheid

De tweede onderzoeksvraag heeft betrekking op de relatie tussen innovatief vermogen en betrokkenheid. Hierbij gaan we ervan uit dat de betrokkenheid van de respondenten op de hoog innovatieve scholen groter zal zijn dan op de laag innovatieve scholen. Het betrokkenheidsmodel biedt uitgangspunten om naar betrokkenheid bij vernieuwingen te kijken. Hiertoe is een aantal meetinstrumenten ontwikkeld, waarvan er in dit onderzoek twee zijn toegepast: het interview 'Gebruiksniveaus van Inno-

vaties' en de vragenlijst 'Intensiteit van Betrokkenheid'. Met deze instrumenten is onderzocht of er in dit onderzoek een verband bestaat tussen innovatief vermogen en de mate van betrokkenheid bij WSNS. In het kader van dit artikel beperken we ons tot de presentatie van enkele resultaten. Voor een meer uitvoerige beschrijving verwijzen we naar Geijsel (1994).

Voor wat betreft het gebruik van een innovatie wordt binnen het betrokkenheidsmodel onderscheid gemaakt tussen acht gebruiksniveaus: van 'geen gebruik' tot 'herziening'. Daar tussenin liggen de niveaus: oriëntatie, mechanisch gebruik (een eerste vormgeving), routine, verfijning (gerichtheid op leerlingen) en integratie (gerichtheid op samenwerking). Tabel 5 geeft de resultaten weer van het interview.

Tabel 5

Verdeling van de verschillende gebruiksniveaus per groep van scholen

N = 64		hoog innovatief	laag innovatief
O/III	oriëntatie	1 (2%)	11 (58%)
III	mechanisch gebruik	10 (22%)	5 (26%)
IV-A	routine	14 (31%)	3 (16%)
IV-B/V	integratie	20 (45%)	0 (0%)

$$\chi^2=36.03, df=3, p=0.00$$

Uit Tabel 5 blijkt dat er een significant verband bestaat tussen het innovatief vermogen van scholen en het gebruiksniveau van WSNS. De hoog innovatieve scholen houden zich meer bezig met de wijze waarop WSNS goed kan worden ingevoerd. Bovendien heeft men meer oog voor wat de vernieuwingen voor leerlingen en collega's kan betekenen. De laag innovatieve scholen houden zich vooral met de eerste fasen van invoering bezig.

Zoals reeds aangegeven is naast de gebruiksniveaus ook de intensiteit van betrokkenheid onderzocht (vragenlijst 'Intensiteit van Betrokkenheid'). Hierbij gaat het om een dimensie die loopt van relatief passief kennis hebben genomen van een vernieuwing, hetgeen een geringe betrokkenheid uitdrukt, tot het opnemen van de vernieuwing in de routine van het lesgeven (Houtveen & Van de Grift, 1993). De maximale waarden op de schaal bij de vragenlijst liggen in principe tussen -3 en +3. Hoe hoger de score op de schaal is, hoe groter de intensiteit van betrokkenheid.

Tabel 6

Gemiddelde scores op de vragenlijst en standaarddeviaties van de twee onderscheiden groepen van scholen

innovatief vermogen	gemiddeld	SD	N
hoog innovatief	-0.65	1.54	45
laag innovatief	-1.65	1.49	19

$t=2.43$, $df=35$, $p=0.02$

In Tabel 6 staan de resultaten weergegeven. Daaruit blijkt dat er een significant verschil is: op de hoog innovatieve scholen is er een grotere intensiteit van betrokkenheid bij WSNS dan op de laag innovatieve scholen. Er dient echter geconstateerd te worden dat deze betrokkenheidsscores relatief laag zijn. Hiervoor zijn enkele verklaringen te geven die te maken hebben met een zekere beperktheid van deze Rasch-schaal bij meerdimensionele vernieuwingen zoals WSNS. In het kader van dit artikel is het niet mogelijk hier verder op in te gaan. Daarvoor verwijzen we naar Geijssel, Van den Berg en Slegers (1995). Ondanks deze relatief lage gemiddelde scores is er bij de hoog innovatieve scholen sprake van een grotere betrokkenheid bij de taakuitvoering en organisatie van WSNS. Er vanuit gaande dat het interview en de vragenlijst beide indicatoren zijn van betrokkenheid kan geconcludeerd worden dat op de hoog innovatieve scholen de betrokkenheid groter is dan op de laag innovatieve scholen in dit onderzoek.

5 Conclusies en discussie

5.1 Conclusies

Wat betreft de eerste onderzoeksvraag laten de resultaten zien dat er duidelijke verschillen zijn tussen hoog en laag innovatieve basisscholen. We vatten deze verschillen per component hieronder kort samen.

De hoog innovatieve scholen blijken de context zoveel mogelijk te gebruiken voor de realisering van hun eigen doelen. Deze scholen percipiëren de beleidsruimte binnen de grenzen die de context stelt. Ze vervullen ook vaak een voortrekkersrol bij grote vernieuwingen als WSNS. De laag innovatieve scholen hebben eerder het gevoel dat de context hen verplicht tot zaken die zij niet willen of waar zij nog niet toe in staat zijn.

De intensiteit van de samenwerking verschilt voor de twee groepen van scholen: op hoog innovatieve scholen wordt meer samengewerkt én op meer verschillende manieren. Enerzijds is de samenwerking klasoverstijgend, mede doordat de mensen een behoefte hebben aan samenwerking. Anderzijds is er sprake van een open sfeer en vertrouwen in elkaar, waardoor goede samenwerking mogelijk is. Op de laag innovatieve scholen daarentegen is gebleken uit veel uitspraken dat niet iedereen even open is; men is bang voor veranderingen en mede daarom wordt de samenwerking liever beperkt tot datgene wat noodzakelijk is voor de voortgang van het lesgeven.

Tussen de leiders van de hoog en de laag innovatieve scholen zijn belangrijke verschillen gevonden. Deze verschillen komen voor een groot deel overeen met de door ons aangegeven aspecten van *transformatief leiderschap* (Leithwood, Tomlinson & Genge, 1995; zie par. 2).

De kenmerken van de *lerende organisatie* zijn samen te vatten onder de noemers groeibehefte en strategievorming. Op hoog innovatieve scholen heeft men een persoonlijke groei-behoefte en wil men zich scholen. Hiertoe worden mogelijkheden geboden, doordat informatie beschikbaar is en nagedacht wordt welke vormen van scholing belangrijk zijn voor de school. Eén van de kenmerken van strategievorming in basisscholen is een middenmanagement waarbij de taken voor iedereen duidelijk zijn. Voorts wordt invloed van leerkrachten bij beleids- en besluitvorming gestimuleerd op de hoog innovatieve school. Strategievorming wordt verder mogelijk gemaakt door de teamvergadering als belangrijkste besluitorgaan aan te wijzen en mensen te stimuleren mee te denken. Lopende zaken zijn bij iedereen bekend, zodat men voldoende op de hoogte is.

Wat betreft de tweede onderzoeksvraag naar de relatie tussen innovatief vermogen en de mate van betrokkenheid bij een vernieuwing kan geconcludeerd worden dat conform de verwachting gebleken is dat op de zes hoog innovatieve scholen er sprake is van een grotere betrokkenheid dan op de vier laag innovatieve scholen: zowel het gebruiksniveau van WSNS als de intensiteit van de betrokkenheid bij WSNS is op

de hoog innovatieve scholen hoger dan op de laag innovatieve scholen.

5.2 Discussie

In het bovenstaande zijn de resultaten van het onderzoek samengevat. Hieronder zullen enkele van deze resultaten diepgaander worden geïnterpreteerd. Het doel van deze *interpretatie* is de oorspronkelijke componenten van het innovatief vermogen nader te specificeren en te nuanceren.

Op basis van de resultaten van dit onderzoek zijn groeibehoeftes en strategievorming gedefinieerd als kenmerken van de component 'lerende organisatie'. In feite wordt het complexe begrip 'lerende organisatie' met deze twee kenmerken niet volledig gemeten. Wat met strategievorming en groeibehoeftes met name wordt geanalyseerd, is de verhouding tussen individuele ontwikkeling en schoolontwikkeling (zie ook Lagerweij & Haak, 1994, pp. 95-113). Karakteristiek aan scholen met een hoog innovatief vermogen is dat zij een afstemming realiseren tussen professionele ontwikkeling en ontwikkeling van de school als geheel. Deze afstemming leidt tot de opvoering van een potentiële kracht van de school om zichzelf te ontwikkelen: de school heeft een zichzelf-ontwikkende kracht. Hiermee sluiten we aan bij Senge (1993; zie Wassink, 1995) die als basis voor een lerende organisatie de individuele motivatie ziet om vanuit een persoonlijke visie zich open en lerend op te stellen, maar wel binnen de context van een team c.q. organisatie (vergelijk ook Swieringa & Wierdsma, 1990, die de verantwoordelijkheid voor het leren in organisaties bij het individu leggen). De resultaten van dit onderzoek lijken met name betrekking te hebben op twee van de vijf door hem geschetste dimensies van de lerende organisatie: persoonlijk meesterschap en teamleren. Zo verwijst groeibehoeftes naar aspecten van persoonlijk meesterschap en strategievorming naar aspecten van teamleren. Voor verder onderzoek naar het innovatief vermogen lijkt het zinvol om uit te gaan van boven beschreven specificering van de component 'lerende organisatie', in termen van 'zichzelf-ontwikkelen-de kracht'.

Bij de definiëring van betrokkenheid (Van den Berg & Vandenberghe, 1995; zie ook par. 2) wordt benadrukt dat deze betrokkenheid een

natuurlijk fenomeen is: het is inherent aan het leerkracht-zijn om de eigen competentie te vergelijken met onderwijsactiviteiten en schoolactiviteiten in het algemeen, waarvan verwacht wordt dat de leerkracht deze op verantwoorde wijze uitvoert. De eigen levensloop, de school waar men werkt, het niveau van de samenwerking etc., zijn allemaal zaken die daarbij een rol spelen. Daarnaast zorgen de voortdurende stroom van vernieuwingen en de daarmee samenhangende permanente druk op de leerkracht ervoor dat betrokkenheid zich als het ware dagelijks gaat manifesteren (Van den Berg & Vandenberghe, 1995). Hieruit blijkt dat betrokkenheid context-bepaald is.

Deze beschrijving maakt duidelijk dat bij het invoeren van vernieuwingen betrokkenheid van leerkrachten een grote rol kan spelen. We herkennen in de theoretische beschrijving van het innovatief vermogen dan ook aspecten van betrokkenheid (zie par. 2). In deze studie zijn aspecten van betrokkenheid gemeten met behulp van twee instrumenten: het interview over de gebruiksniveaus en de vragenlijst naar de intensiteit van betrokkenheid. Echter, uit de interviews naar het innovatief vermogen komen ook (soms impliciet, soms expliciet) vragen van betrokkenen naar voren die een onzekerheid signaleren en tevens wijzen op een bereidheid tot innoveren. Zo verwoordden de geïnterviewden een duidelijke wens tot individuele ondersteuning, geven ze aan behoefte te hebben aan scholing en aan klasoverstijgende samenwerking, wijzen ze op de wenselijkheid van een open sfeer voor professionele ontwikkeling. Deze wensen, behoeften en vragen verwijzen in feite naar aspecten van betrokkenheid, zoals gemeten met de afzonderlijke betrokkenheidsinstrumenten.

Zo geïnterpreteerd lijkt het interview naar het innovatief vermogen tevens aspecten van betrokkenheid te meten. Een verdere conceptuele doordinking van de relatie tussen betrokkenheid en innovatief vermogen is daarom gewenst.

Met deze interpretaties van de resultaten hopen we het verder onderzoek naar het innovatief vermogen vorm te geven en een basis te leggen voor de ontwikkeling van een vragenlijst naar het innovatief vermogen.

- Bakx, E., & Eerden, E. van der (1993). *Onderzoek naar het innovatief vermogen van scholen in het voortgezet onderwijs* (Doctoraalscriptie). 's-Hertogenbosch: Katholiek Pedagogisch Centrum/ Nijmegen: Katholieke Universiteit, Vakgroep Onderwijskunde.
- Beek, A. ter (1992). *Betrokkenheidsinstrumenten vergeleken* (Doctoraalscriptie). 's-Hertogenbosch: Katholiek Pedagogisch Centrum/ Nijmegen: Katholieke Universiteit, Vakgroep Onderwijskunde.
- Berg, R. van den (1990). *Transformatief vernieuwingsbeleid* (Inaugurale rede). Deventer: Van Loghum Slaterus.
- Berg, R. van den (1992). Transformational education policy: Schools as strategic organization. *Knowledge: Creation, Diffusion, Utilization*, 13, 440-459.
- Berg, R. van den (1993). The concerns-based adoption model in the Netherlands, Flanders and the United Kingdom: state of the art and perspective studies. *Educational Evaluation*, 19, 51-63.
- Berg, R. van den, Slegers, P., Bakx, E., & Eerden, E. van der (1994). Het innovatief vermogen van scholen in het voortgezet onderwijs: Een kwalitatief onderzoek. *Pedagogische Studiën*, 71, 402-419.
- Berg, R. van den, & Vandenbergh, R. (1981). *Onderwijsvernieuwing in verschuivend perspectief*. Tilburg: Zwijzen.
- Berg, R. van den, & Vandenbergh, R. (1984). *Groot-schaligheid in de onderwijsvernieuwing*. Tilburg: Zwijzen.
- Berg, R. van den, & Vandenbergh, R. (1988). *Onderwijsvernieuwing op een keerpunt*. Tilburg: Zwijzen.
- Berg, R. van den, & Vandenbergh, R. (Eds.) (1995). *Wegen van betrokkenheid: Reflecties op onderwijsvernieuwing*. Tilburg: Zwijzen.
- Bomers, G. B. J. (1989). *De lerende organisatie*. Nijmegen, opening van het academisch jaar.
- Cobbenhage, J. W. C. M. (1990). Innoveren: strategieën en modellen. In J. F. den Hertog & F. M. van Eijnatten (Eds.), *Management van technologische vernieuwing* (pp. 8-40). Assen/Maastricht: Van Gorcum.
- Fullan, M. G., & Stiegelbauer, S. (1991). *The new meaning of educational change*. New York/ Toronto: Teachers College Press.
- Geijssels, F. (1994). *Het innovatief vermogen van scholen in het basisonderwijs* (Doctoraalscriptie). 's-Hertogenbosch: Katholiek Pedagogisch Centrum/ Nijmegen: Katholieke Universiteit, Vakgroep Onderwijskunde.
- Geijssels, F., Berg, R. van den, & Slegers, P. (1995). Innovatief vermogen van scholen en betrokkenheid. In R. van den Berg & R. Vandenbergh, (Eds.), *Wegen van betrokkenheid: Reflecties op onderwijsvernieuwing*. Tilburg: Zwijzen.
- Gennip, J. van (1991). *Veranderingscapaciteiten van basisscholen* (Academisch proefschrift). Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen (ITS)/ Heerlen: Open Universiteit.
- Gerards, J. (1993). *Essentiële vernieuwingsbehoeften van docenten* (Concept-versie van doctoraalscriptie). 's-Hertogenbosch: ongepubliceerd.
- Houtveen, Th., & Grift, W. van de (1993). *De intensiteit van betrokkenheid van onderwijsgevers bij het werken aan onderwijsvernieuwingen*. Utrecht: Interdisciplinair Sociaal Wetenschappelijk Onderzoeksinstituut (ISOR).
- Huberman, M. (1993). The model of the independent artisan in teachers' professional relations. In J. W. Little & M. W. McLaughlin (Eds.), *Teachers work, individuals, colleagues and contexts* (pp. 11-50). New York: Teachers College Press.
- Kuijlaars, H. (1992). 'Betrokkenheid' treffend verwoord: Een kwalitatieve invulling van het interviewinstrumentarium 'Gebruiksniveaus van Innovatie' (Doctoraalscriptie). 's-Hertogenbosch: Katholiek Pedagogisch Centrum/Nijmegen: Katholieke Universiteit, Vakgroep Onderwijskunde.
- Lagerweij, N. A. J. (1993). Implementatie van basisvorming: de start van een dieptestudie. *Pedagogische Studiën*, 70, 162-180.
- Lagerweij, N. A. J., & Haak, E. (1994). *Eerst goed kijken.... De dynamiek van scholen-ontwikkeling*. Leuven/Apeldoorn: Garant.
- Leithwood, K. A. (1994). *The contributions of transformational leadership to organizational learning in schools*. Paper presented at the annual meeting of the American Educational Research Association (AERA), New Orleans.
- Leithwood, K., Tomlinson, D., & Genge, M. (1995). Transformational school leadership. In K. A. Leithwood (Ed.), *International handbook on educational leadership*. Kluwer Press. (In druk)
- Louis, K. S. (1994). Beyond 'managed change': rethinking how schools improve. *School Effectiveness and School Improvement*, 5(1), 2-24.

Auteurs

F. Geijssels is als assistent in opleiding verbonden aan de Vakgroep Onderwijskunde van de Katholieke Universiteit te Nijmegen.

R. van den Berg is als bijzonder hoogleraar verbonden aan de Vakgroep Onderwijskunde van de Katholieke Universiteit te Nijmegen en geeft leiding aan het onderzoeksteam van de KOMPACTGROEP/KPC te 's-Hertogenbosch.

P. Sleegers is als universitair docent verbonden aan het Universitair Instituut voor de Lerarenopleiding (Unilo) van de Katholieke Universiteit Nijmegen.

Adres: Katholieke Universiteit Nijmegen, Vakgroep Onderwijskunde, Postbus 9104, 6500 HE Nijmegen

Abstract

The innovative capacity of schools in primary education: a second qualitative study

F. Geijssels, R. van den Berg & P. Sleegers. *Pedagogische Studiën*, 73, 42-55.

Schools are increasingly responsible for their own futures. In this light, it may be fruitful to examine the capacity of these organizations for innovation, in particular their capacity to realize large-scale innovations. With this goal in mind, a strategy for the study of the innovative capacity of schools was developed.

In a previous article for this journal, a preliminary study of the innovative capacity of secondary school was reported. The results of this research were then used to create a conceptual model of the innovative capacity of schools.

In the present study, the innovative capacity of primary schools in The Netherlands was examined. The purpose of the present study was to validate the model developed in the previous study and go one step further by examining the possible relation between the innovative capacity of schools and the intensity of concerns of teachers and school leaders.

A number of high-innovation and low-innovation schools was selected. A questionnaire was administered and two interviews were undertaken with the teachers, internal supporters, and school leaders in these primary schools ($N = 64$). The results of the questionnaire and the interviews provide a better understanding of the innovative capacity of schools.

- Marx, E. C. H., & Vries, O. A. de (1995). *School en klas*. In J. Lowyck & N. Verloop (Eds.), *Onderwijskunde: Een kennisbasis voor professionals*. Groningen: Wolters-Noordhoff.
- Mintzberg, H. (1991). *Mintzberg over management*. Amsterdam/Antwerpen: Uitgeverij Veen.
- Morgan, G. (1986). *Images of organizations*. London: SAGE Publications.
- Rolff, H.-G. (1993). *Wandel durch Selbstorganisation: Theoretische Grundlagen und praktische Hinweise für eine bessere Schule*. Weinheim/München: Juventa Verlag.
- Salisbury, D. F., & Conner, D. R. (1994). How to succeed as a manager of an educational change project. *Educational Technology*, 34(6), 12-19.
- Senge, P. M. (1993). *The fifth discipline: the art and practice of the learning organisation*. New York: Doubleday, Currency.
- Sleegers, P. J. C. (1991). *School en beleidsvoering: Een onderzoek naar de relatie tussen het beleidsvoerende vermogen van scholen en het benutten van de beleidsruimte door scholen* (Academisch proefschrift). Nijmegen: Katholieke Universiteit.
- Sleegers, P. J. C., & Wesselingh, A. (1993). Decentralization in education: A dutch study. *International Studies in Sociology of Education*, 3(1), 49-67.
- Smaling, A. (1992). Varieties of methodological intersubjectivity: Relations with qualitative and quantitative research and with objectivity. *Quality & Quantity*, 26, 169-180.
- Sundbo, J. (1992). The tied entrepreneur: On the theory and practice of institutionalisation of creativity and innovation in service firms. *Creativity and Innovation management*, 1(3), 109-120.
- Swieringa J., & Wierdsma, A. F. M. (1990). *Op weg naar een lerende organisatie*. Groningen: Wolters-Noordhoff.
- Vandenberghe, R. (1993). Grootschalig vernieuwen: Samenspel tussen beleid, begeleiding en scholen. *School & Begeleiding*, 39, 18-21.
- Wassink, H. (1995). *Aspecten van innovatief vermogen in twee verpleeghuizen* (Doctoraalscriptie). Nijmegen: Katholieke Universiteit, Vakgroep Onderwijskunde.
- Welten, V., & Janssen, J. (1993). *Teksttabel: een gecomputeriseerde methode voor het analyseren van tekstuele data*. Nijmegen: Katholieke Universiteit, Vakgroep Cultuur- en Godsdienstpsychologie.

Onderzoek beoordeeld

Door de redactie is mij gevraagd als voorzitter van de internationale onderzoeksbeoordelingscommissie pedagogische en onderwijskundige wetenschappen mijn persoonlijke indrukken te geven over het beoordelingsproces en de resultaten daarvan. Omdat mij uit reacties van faculteiten op het concept-rapport is gebleken dat er ondanks de gedegen vóór-informatie door de VSNU in het Protocol 1994 nogal wat misverstanden heersten omtrent die onderzoeksbeoordeling, lijkt het mij noodzakelijk aan mijn persoonlijke indrukken enige feitelijke informatie over het beoordelingsproces te doen voorafgaan.

Vooraf aan de werkwijze van de commissie en aan de wijze waarop de 'ratings' tot stand kwamen wil ik enige aandacht schenken, omdat daarover grote onduidelijkheid bleek te bestaan.

Met de commissie acht ik mij echter niet geroepen om op verzoek van enkele faculteiten het verschil aan te geven tussen tellingen ('scores', bijv. van productiecijfers) en intersubjectieve 'peer ratings' (bijv. van de produktie). De commissie meent dit verschil bekend te mogen veronderstellen.

Wie trouwens meent dat het merendeel van de hier ten overvloede gegeven informatie in het rapport valt te lezen, moet ik daarin gelijk geven, al is het mij niet ontgaan dat niet iedereen tot de goede lezers blijkt te behoren. In een enkel geval verstrekt mijn informatie, denk ik, wat meer inzicht dan het 'final report'.

De werkwijze van de commissie was als volgt. Het eerste beoordelingsmoment was de schriftelijke beoordeling volgens een te voren vastgesteld 'assessment form' van de wetenschappelijke rapporten van de programma's der verschillende faculteiten door de leden van de commissie. Hun was door de voorzitter gevraagd de rapporten in hun geheel te bestuderen, maar in het bijzonder aandacht te schenken aan de programma's die hun als eerste of tweede reviewer door de voorzitter waren toegewezen. De voorzitter zelf trad slechts als

tweede reviewer op. Uiteraard werd bij deze toewijzing zoveel mogelijk rekening gehouden met de speciale expertise van de reviewers.

In de maartsessie van de commissie in Amersfoort speelde zich een eerste discussie af tussen de eerste en de tweede reviewers naar aanleiding van hun huiswerk.

Nadat zij overeenkomst hadden bereikt over hun ratings aan de hand van het Protocol 1994 had een plenaire discussie plaats in de gehele commissie. Na een stevige uitwisseling van standpunten en argumentatie daaromtrent leidde deze tot een door de voltallige commissie gedeelde voorlopige rating.

In de sessie van eind april, eveneens te Amersfoort, werden alle programmaleiders, al of niet van anderen vergezeld, geïnterviewd naar aanleiding van door de commissie toegezonden vragen en opmerkingen. Daarna werden door de commissie de voorlopige ratings opnieuw bezien en indien nodig bijgesteld in een soortgelijke procedure als in maart, eerst door de beide reviewers, daarna in een discussie in de voltallige commissie.

De ratings werden nu door de commissie als definitief beschouwd, eventueel nog bij te stellen na toezending van het concept-rapport aan de faculteiten en de reacties daarop. In een enkel geval heeft die bijstelling plaats gevonden.

De overige dagen werden door de commissie gebruikt om een eerste 'draft' van het rapport op te stellen en te becommentariëren.

De leden gingen naar huis met de opdracht een meer definitief concept aan te leveren, waarbij ook aan voorzitter en secretaris taken werden toegewezen.

De reacties van de faculteiten op het verzonden concept-rapport werden door de leden bestudeerd, waarna het rapport definitief werd bijgewerkt en vastgesteld. De brieven aan de faculteiten werden na overleg beantwoord.

Hoewel het concept-rapport direct na het verstrijken van de in het Protocol genoemde tijdsbestek werd toegezonden viel dit ongelukkigwijze samen met het begin van de academische vakantie, zodat aan de faculteiten tot half september de tijd moest worden gegund om te reageren. Pas daarna kon de commissie die antwoorden bestuderen.

Op enkele reacties van faculteiten wordt in het volgend deel van deze informatie of in mijn persoonlijk commentaar teruggekomen.

Omdat ondanks de beschrijving in het Protocol 1994 en in het concept-rapport misverstanden bleven bestaan omtrent de wijze waarop ratings tot stand kwamen, wil ik daarop nog summier ingaan.

Als criteria van *kwaliteit* golden voor de commissie: de key-publications, de publikaties in vooral Engelstalige, buitenlandse high standard tijdschriften en andere gerefereerde buitenlandse publikaties (boeken), de publikaties in andere, vooral Engelstalige tijdschriften, de publikaties in gerefereerde Nederlandse en eventueel Franse en Duitse tijdschriften en boeken.

Aangezien er een hoge correlatie blijkt te bestaan tussen het aantal artikelen in buitenlandse tijdschriften in het algemeen en het aantal artikelen in buitenlandse 'high impact' tijdschriften is een onderscheid tussen 'high impact' en 'low impact' tijdschriften, op de manier waarop dat door de commissie psychologie is gehanteerd, niet toegepast.

Verder werden in de kwaliteitsbeoordeling betrokken: (inter)nationale awards, lidmaatschappen van internationale wetenschappelijke gezelschappen, redacteurschappen van internationale en van gerefereerde Nederlands-talige tijdschriften. Een enkele programmeerder was van mening dat de commissie te veel waarde hechtte aan de internationale publikaties. De commissie vindt dat het internationale karakter van wetenschapsbeoefening dat rechtvaardigt.

Wetenschapsparadigma's speelden alleen een rol wanneer de commissie die als gedateerd beschouwde.

Eén faculteit maakte daar in haar reactie bezwaar tegen, omdat zij dit geen aspect van kwaliteit van publikaties achtte. De commissie had daarvoor wel een zeker begrip om een zo dadelijk uiteen te zetten reden, maar zij hield toch vast aan haar standpunt. Als aspect van wetenschappelijke kwaliteit had zij immers als eerste punt genoemd "the originality of approaches and ideas in tackling scientific problems", evenals andere commissies, bijv. die van "Electric Engineering" dat ook hebben gedaan. Tot die 'original approaches' rekende zij ook het paradigma. Het 'zekere begrip' van de commissie sproot voort uit de gedachte dat men met een zekere relativering over deze kwestie behoort te denken. Want al is bijv. in de onder-

wijskunde het behavioristische paradigma volstrekt gedateerd, het thans in zwang zijnde constructivistisch paradigma is in wezen reeds aan te wijzen in de 'discovery learning'-benadering van Piaget en Bruner, maar toch kennelijk niet gedateerd geacht. Die relativering leidde tot opname van een appendix van betrokken faculteit.

De beoordeling van de *productiviteit* heeft volgens de commissie een kwalitatieve en een kwantitatieve, numerieke kant. De commissie sprak dus niet louter op basis van numerieke gegevens een oordeel uit. Om die reden lette zij ook minder op het volume dan op de kwaliteit van het produkt, al werd een artikel van 2 of 3 pagina's veelal niet hoog gewaardeerd. De kwaliteit van de produktie beoordeelde de commissie op basis van de '1- and 2-rule', een internationaal erkende regel, waarin het cijfer 1 aangeeft het aantal publikaties in gerefereerde internationale tijdschriften of handboeken, en het cijfer 2 het aantal andere internationale publikaties per fte per jaar. In deze zin is toch uitdrukking gegeven aan het onderscheid tussen high impact publikaties en andere.

De kwantitatieve wetenschappelijke output (scoring), vooral in de internationale pers speelde eveneens een rol. Vakpublicaties werden niet in de beoordeling betrokken. Uiteraard heeft de commissie niet van elke publikatie aangegeven of zij dit een wetenschappelijke publikatie of een vakpublicatie vond. Dat is ook volstrekt onnodig na wat hiervoor over haar beoordeling van publikaties is opgemerkt. Zij heeft wel voorbeelden gegeven van wat zij als vakpublicaties beschouwde. Als bij wijze van zulk een voorbeeld artikelen in *Onderwijskundig Lexicon*, etc. werden opgevoerd als wetenschappelijke publikaties heeft de commissie die niet als zodanig gewaardeerd. Het voorwoord van dit lexicon, editie II geeft zelf duidelijk aan dat het daar om vakpublicaties gaat.

Het *relevantie*-aspect betrof de bijdrage aan de uitbreiding van wetenschappelijke kennis, maar ook het belang van het programma voor de praktijk van opvoeding en onderwijs en voor het beleid.

Het oordeel over de *viabiliteit* betrof de toekomstige taak die voor het programma gezien werd, uitgaande van de ontwikkeling in de tijd en rekening houdend met de persoonlijke en materiële middelen die ter beschikking stonden.

Kom ik nu tot mijn *persoonlijk commentaar*, dan wijs ik eerst op het door mij duidelijk ervaren belang van een evaluatie van onze universitaire onderzoeksprogramma's door een internationale commissie, die ons deels nog wat provincialistisch pedagogische wereldje een spiegel voorhoudt. Dat het niet voor iedereen prettig is het eigen spiegelbeeld te zien, moet men denk ik voor lief nemen, al mag men best op onderdelen met de commissie van mening verschillen.

Wat de beoordeling voor de commissie niet eenvoudig maakte is de grote diversiteit van subdisciplines: wijsgerige en historische pedagogiek, empirisch pedagogische studies, studies in jeugdkunde en in volwasseneneducatie, orthopedagogiek met een brede mix van leren en gedragsstoornissen van kinderen en jeugdigen, onderwijskunde op micro-, meso- en macroniveau, onderwijstechnologie, statistische en mathematische methodieken, voor welke laatste de commissie externe adviseurs om een oordeel heeft gevraagd.

De discussie over de programma's in al die subdisciplines werd wel in de gehele commissie gevoerd, maar de experts hadden uiteraard een belangrijke invloed in die discussie.

Een punt van zorg lijkt mij de verhouding tussen de fundamentele research en de theoretische basis daarvan en de toegepaste research, welke laatste nogal oververtegenwoordigd was. Dat is wellicht begrijpelijk in wetenschappen die zo sterk op het handelen betrokken zijn als pedagogiek en onderwijskunde, maar dat neemt niet weg dat de basis van fundamentele research ook in deze wetenschappen onmisbaar is en m.i. meer nadruk moet krijgen. In de internationale vergelijking draagt het Nederlands onderzoek wat sterk een toegepast karakter. Was er daarom zo weinig excellent onderzoek te bespeuren?

Wellicht vallen de oordelen van de commissie mede daardoor wat lager uit dan menig onderzoeker blijkens de reacties gehoopt of verwacht had. Het Nederlands onderzoek is volgens internationale maatstaven in het algemeen van een bevredigend niveau ('satisfactory'), maar niet meer dan dat.

Een enkele faculteit wist dat zo moeilijk te verwerken dat zij het rekenwerk van de commissie volgens een andere methode dan die van de commissie heeft overgedaan en daardoor tot

andere conclusies is gekomen dan de internationale commissie. Deze heeft van die bewering met belangstelling kennis genomen zonder zich echter van de eigen koers te laten afleiden.

Er waren zelfs onderzoekers die zich, buiten medeweten van de voorzitter van de commissie, hebben gewend tot hun bekende individuele leden van de commissie om die tot andere gedachten te brengen, hetgeen de commissie uiteraard niet bijzonder apprecieerde.

Weer andere researchers vergeleken hun scores en ratings met die van andere universiteiten en suggereerden dat zij in vergelijking met die anderen onrechtvaardig behandeld waren. Uit deze voorbeelden blijkt duidelijk, hoe belangrijk men een evaluatie als deze vindt en hoe moeilijk men het vindt de uitkomsten daarvan te accepteren als die ongunstiger uitvallen dan men verwacht had.

Dat zij groot belang hechtten aan het oordeel van de commissie blijkt ook uit het feit dat onderzoekers van een der universiteiten niet konden wachten op het definitieve rapport, alvorens zichzelf op de borst te kloppen, en hebben gelect naar de eigen universitaire pers die hun succes voortijdig publiceerde.

De gegeven voorbeelden vormen uitzonderingen. De meeste faculteiten stelden zich, na op feitelijke onjuistheden in het concept-rapport gewezen te hebben, positief op ten aanzien van het werk van de commissie, ook wanneer de oordelen voor de eigen faculteit minder gunstig waren. Het rapport geeft immers duidelijk aan waar de sterkten en de zwakten van programma's liggen en doet in diverse gevallen suggesties voor verbetering die de viabiliteit van de programma's kunnen versterken en de kwaliteit kunnen verbeteren. Minder in de breedte, meer zich beperken tot in de diepte werken is meermalen het advies.

Ook is aansluiting bij onderzoeksscholen en versterking van de invloed daarvan een advies dat werd verstrekt.

Wat zou de VSNU volgens mijn perceptie kunnen doen ter verbetering van de procedure?

Aan het voorbereidend werk van de VSNU (zie het Protocol 1994) en de ondersteuning van het werk van de commissie kan niet veel worden verbeterd. (Dat het concept-rapport

precies volgens de protocollaire regels bij de faculteiten arriveerde op het moment dat de academische vakantie begon, is in het vervolg door planning te voorkomen).

Dat onze commissie het beter vond met een 10-punts dan met een 5-punts ratingschaal te werken is een kleinigheid en wordt wellicht ook door andere commissies niet gedeeld.

Een andere kleinigheid is dat het voor de commissie gemakkelijk zou zijn, en door de VSNU bevorderd kan worden, dat alle faculteiten internationale en nationale publikaties in aparte lijsten zouden vermelden.

Overweging verdient ook om aio's niet meer mee te tellen bij het wp, daar de meeste aio's niet vóór hun promotie publiceren. Sommige faculteiten hebben daar terecht op aangedrongen.

De verbetering van de procedure moet echter vooral van andere zijden komen. Zo is mij van verscheidene kanten verteld dat leden van de commissie onbekenden voor betrokkenen waren. Nu kan dat het gevolg zijn van het gering aantal internationale contacten van een aantal onderzoekers. De curricula vitae van de gewone leden laten zien dat het om internationaal bekende persoonlijkheden gaat.

Toch zou de vraag gesteld kunnen worden of het samenspel tussen de disciplineoverlegorganen en de KNAW van zodanige aard is geweest dat die organen zich voldoende in de keuze van de KNAW konden herkennen.

Van groter belang is een tweede punt: er valt echt wat te doen van de zijde van de faculteiten. Wie de divergentie van de oordelen van de commissie over verschillende programma's binnen vrijwel elk van de faculteiten overziet, kan niet anders dan tot de conclusie komen dat een aantal programma's niet ter beoordeling ingediend had moeten worden. De faculteiten hebben blijkbaar geen voldoende invloed om op de rem te trappen als zo'n zwak programma toch voor beoordeling wordt voorgedragen.

Nu er onderzoeksscholen bestaan in Nederland, waaraan de universiteiten, de minister, maar ook de commissie grote waarde hechten, zou het dan niet beter zijn dat door die onderzoeksscholen een vóórselectie werd gemaakt voor de indiening van internationaal te beoordelen programma's? Een beroepscommissie kan dan eventueel tevoren een uitspraak doen

als faculteiten problemen hebben met de keuze van de onderzoeksscholen. Maar de internationale beoordeling moet daaraan volgend wel blijven bestaan.

Een laatste persoonlijke ervaring, maar die ook door de meeste leden van de commissie wordt gedeeld is dat het werken in dit internationaal verband inspirerend, aangenaam, maar volgens de aangegeven procedure ook zeer vermoeiend is. Wellicht moet men in het vervolg een jongere, in het buitenland werkende voorzitter voor zulk een commissie benoemen.

A. M. P. Knoers

Zesde EARLI-Conferentie Nijmegen, 26-31 augustus 1995

Algemene indrukken (J.M. Pieters, Universiteit Twente)

Tien jaar na de eerste conferentie in Leuven had de EARLI in augustus jl. haar tenten opgeslagen in Nijmegen. Dat de EARLI in de tien jaar van haar bestaan uitgegroeid is tot een volwaardige Europese onderzoekersorganisatie moge blijken uit het feit dat niet alleen collega's uit Europese landen aan de conferentie deelnamen, maar dat nogal wat collega's uit de Verenigde Staten de moeite hadden genomen naar Europa af te reizen. De organisatie had als centrale thema *Learning to learn and teaching how to learn* gekozen. Een niet zo verrassende keuze als men bedenkt dat de lokale organisatoren zich in belangrijke mate op dit terrein hebben gemanifesteerd. Dit thema kwam overduidelijk aan de orde in de tien onderzoekdomeinen die voor de conferentie als organisatiekader dienden: cognitie en meta-cognitie; ontwikkeling; motivationele, sociale en affectieve processen; onderwijzen en instructietheorie; onderwijzen, leren en technologie, en het ontwerpen van instructie; specifieke domeinen; sociale en culturele context en de huiselijke omgeving; volwasseneneducatie en bedrijfsopleidingen; hoger onderwijs; en methodologie en peiling. Het centrale thema klonk ook door in de *invited addresses* van gerenommeerde sprekers uit Europa, Japan, Verenigde Staten en Canada. De organisatie had, zo vermeldde het programmaboek, meer dan 350 presentaties, waarvan ongeveer 200 individuele posters, 76 verschillende symposia, ongeveer 14 demonstraties en ronde-tafel discussies, en meer dan vijf workshops en panel discussies verzameld.

In deze terugblik op de EARLI-conferentie in Nijmegen komen onderzoekers aan bod die ieder vanuit zijn eigen onderzoeksperspectief een impressie geeft van de conferentie. In elk

van de te bespreken onderwerpen komt het centrale thema aan de orde. Een opvallende constante factor in elke bijdrage is het vigerende constructivisme, door de auteurs, in overeenstemming met de uitgangspunten van het constructivisme, op geheel eigen wijze geïnterpreteerd. Deze kroniek pretendeert geen alomvattend overzicht van de EARLI-conferentie te verstrekken. In plaats daarvan is gekozen voor het belichten van een aantal thema's en domeinen. Voor nadere informatie over de hier behandelde en niet-behandelde thema's en domeinen verwijs ik graag naar het Abstract-boek van de conferentie (Aarnoutse, C., Jong, F. de, Lodewijks, H., Simons, R.J., Aalsvoort, D. van der, (Eds.) (1995). *Abstracts 6th EARLI-conference*. Tilburg: MesoConsult).

Meta-cognitie (M.V.J. Veenman, Rijksuniversiteit Leiden)

Meta-cognitie was een relevant subthema van het EARLI-congres over 'leren te leren'. Eén key-note, twee symposia en drie poster-sessies waren geheel aan dit onderwerp gewijd. Bovendien werd in zeventien verspreide papers en posters aandacht aan het onderwerp besteed. Opvallend was dat deze bijdragen nadrukkelijk twee stromingen in de onderzoeksliteratuur vertegenwoordigden.

Enerzijds is er onderzoek dat is gericht op de rol van meta-cognitieve kennis, dat wil zeggen de (declaratieve) kennis die iemand heeft over het eigen cognitieve systeem. Anderzijds is er onderzoek dat zich richt op meta-cognitieve vaardigheden, dat wil zeggen de (procedurele) kennis over hoe men het eigen probleemoplossen leergedrag daadwerkelijk stuurt en controleert. Dergelijke sturingsactiviteiten als de oriëntatie op de taak, het plannen van activiteiten, monitoring en evaluatie worden ook wel met de term zelfregulatie aangeduid.

Beide onderzoeksstromingen verschillen zowel in studie-object als in onderzoeksmethodiek. De ene groep onderzoekers is vooral geïnteresseerd hoe de verwerving van meta-cognitieve kennis wordt beïnvloed door motivationele en attributie-processen. Meta-cognitieve kennis wordt doorgaans gemeten met behulp van vragenlijsten en interviews. Op dit EARLI-congres waren met name Boekaerts e.a., Demetriou en Efklides, Lompscher e.a., en

Pintrich en McKeachie vertegenwoordigers van deze stroming. Een algemene bevinding uit deze bijdragen was dat meta-cognitieve kennis maar matig correleert met leerprestaties. Bovendien vonden Lompscher e.a. zeer lage correlaties tussen meta-cognitieve kennis, gemeten met een vragenlijst, en het feitelijk vertonen van meta-cognitieve activiteit tijdens het uitvoeren van een leertaak. Kennis van het eigen cognitief functioneren impliceert derhalve niet het daadwerkelijk sturen en reguleren van dit functioneren.

De tweede groep onderzoekers richt zich op de relatie tussen meta-cognitieve vaardigheden en leerprestaties, en op de vraag wat de voorwaarden zijn voor een effectieve training van meta-cognitieve vaardigheden. Technieken voor het tijdens de taakuitvoering meten van meta-cognitieve vaardigheden zijn veelal systematische observatie en protocol-analyse. Bijdragen op dit terrein werden onder meer geleverd door Salomon, Tobias, Veenman, Volet, en Winne. Tobias presenteerde een nieuw instrument voor het meten van meta-cognitieve kennis en vaardigheden, waarbij een zelfbeoordelingslijst werd ingevuld tijdens de uitvoering van eenvoudige cognitieve taken. Deze, aan taakuitvoering gekoppelde zelfbeoordeling bleek een relevante predictor voor latere schoolresultaten bij diverse vakken. In een symposium over het nut van het trainen van meta-cognitieve vaardigheden stelde Volet dat een meta-focus op training (d.w.z. de inhoud van de meta-cognitieve training) zou moeten worden gecombineerd met een meta-vorm van instructie (d.w.z. de instructie-methodiek). Bekende voorbeelden van meta-vormen van instructie zijn Brown en Palincsar's *scaffolding* en *reciprocal teaching*. Salomon voegde daaraan toe dat niet alleen een beoogde meta-cognitieve vaardigheid moet worden getraind maar ook de strategische of conditionele kennis over hoe en wanneer de vaardigheid kan worden toegepast. Veenman stelde dat een dergelijke inzet van methodieken afhankelijk is van het doel van meta-cognitieve training. Is het doel louter verbetering van de overdracht van domeinspecifieke kennis dan kan men volstaan met meta-cognitieve instructie op het juiste moment tijdens de taakuitvoering. Beoogt men echter een algemene transfer van meta-cognitieve vaardigheden voorbij de trainingssituatie,

dan zullen zelfs de principes van Volet en Salomon niet toereikend zijn. Dat vereist een ander schoolsysteem waarin meta-cognitieve vaardigheden consequent én multidisciplinair worden getraind. Verder pleitte Veenman voor onderzoek naar de ontbrekende schakel tussen meta-cognitieve kennis en vaardigheden. Een interessante aanzet daartoe werd gegeven door Winne in zijn lezing over de verwerving van zelf-regulatieve strategieën. Bij het uitvoeren van een bepaalde taak selecteert men een 'tactic', een algemene strategie voor informatie-verwerking (Marzano), uit een repertoire van beschikbare tactics. Een tactic wordt geselecteerd op grond van de kennis die men heeft over de aard van de taak, het gestelde doel, de effectiviteit van beschikbare tactics, en de condities waaronder een tactic effectief kan worden toegepast (Salomon). Deze (meta-cognitieve) kennis wordt opgebouwd en verbeterd tijdens successievelijke leerervaringen volgens een test-operate-test-exit (TOTE) principe. Volgens Winne worden individuele verschillen in zelfregulatie derhalve veroorzaakt door de mate waarin men in staat is zelfstandig goede leerervaringen te creëren en uit deze leerervaringen meta-cognitieve kennis te abstraheren.

Tot slot volgt een korte bespreking van het symposium *Teaching intelligence* met bijdragen van Klauer, Vizcarro e.a., en Adey. Opvallend aan intelligentie-trainingsprogramma's, zoals dat van Klauer, is dat zij een sterke meta-cognitieve trainingscomponent hebben. Eventueel succes van zo'n training (op een analoge intelligentietaak of een leertaak) kan derhalve aan zowel cognitieve als meta-cognitieve factoren worden toegeschreven. Werk van Veenman en van Minnaert en Janssen (elders op deze conferentie) heeft laten zien dat meta-cognitieve vaardigheid niet mag worden vereenzelvigd met gemeten intelligentie. Meta-cognitieve vaardigheden kunnen onder bepaalde voorwaarden worden getraind; van intelligentie opgevat als het repertoire van cognitieve vaardigheden valt dat nog te bezien. Elshout heeft de training van intelligentie ooit het opwrijven van de thermometer genoemd: voor algemene transfer van deze vaardigheden, zoals Adey voorstaat, bestaat immers nauwelijks evidentie.

De onderwerpen onderwijs en instructie waren op deze zesde EARLI-conferentie nauwer op elkaar betrokken dan ooit. Waar nog in vorige conferenties sprake was van een apart domein, lerarenopleiding, in Nijmegen werd dit domein behandeld in combinatie met andere, te weten: *teaching; instructional theory en teaching; learning and technology*.

Nu was het domein van de leraar c.q. de lerarenopleiding nooit prominent vertegenwoordigd op de EARLI-conferenties, ook dit jaar was dit niet het geval. Er waren welgeteld één *invited address* (waarover meer), slechts drie symposia (van de 76) en 12 *structured posters* (van de 200), die in het bijzonder gericht waren op de leraar en de lerarenopleiding. Daar staat tegenover dat, meer dan in andere jaren, het label *Teaching* te vinden was in presentaties die vielen binnen andere domeinen. Ik wijt dit aan de 'constructivistische' wending in de instructiepsychologie die meer aandacht heeft doen ontstaan voor 'sturingsmechanismen voor verandering' van leerprocessen en bijgevolg voor de 'methodische' en contextuele kant van instructie (zodat men uitkomt bij de onderwijzen-kant van instructie).

Als ik mij beperk tot een strikte aandacht voor onderwijzen en degene die onderwijst, dat is de leraar, viel het mij des te meer op dat de genode inleider H. Francis zich bezighield met een nauwkeurige terreinafbakening van psychologie en lerarenopleiding. Zij zocht een nieuwe rol voor de discipline psychologie in de lerarenopleiding. Zij wees op het, al lang niet meer Angelsaksische, verschijnsel dat de psychologie - in ons land mag daar ook de onderwijskunde aan worden toegevoegd - veel invloed en terrein heeft verloren in de lerarenopleiding. Zij pleitte voor een heroriëntatie in de psychologie waarbij deze (en de onderwijskunde, zou ik aanvullen) meer dienstverlenend, meer vanuit een integrale benadering van het opleidingscurriculum zal opereren. Geen nieuw geluid en geen nieuwe remedie, maar het moest kennelijk nogmaals gezegd worden.

De poster sessies bevatten zeer uiteenlopende onderwerpen en kenden meestal een flinke bijdrage vanuit Nederlands/Belgisch onderzoek. Omdat dit onderzoek zich in dit tijd-

schrift wel weet aan te dienen, geef ik op deze plaats er de voorkeur aan aandacht op ander werk te richten, zoals bijvoorbeeld dat van Mevorach naar mentale modellen van leraren en hun leerlingen. Dit onderzoek toont het belang aan van diagnostische kennis bij leraren over hun leerlingen. Deze kennis is bepalend voor het al dan niet ingaan door de leraar op misvattingen van leerlingen. Verder werden enkele inventariserende studies gepresenteerd over leren in de opleiding door studenten en de noodzaak van nieuwe aanpakken voor het leren onderwijzen. Opvallend was dat in de posters zo weinig aan resultaten werd gepresenteerd. Men presenteerde (en beschikte kennelijk enkel over) een conceptuele voorbereiding en een methodische opzet van onderzoek.

De symposia kenden duidelijk meer structuur. Bennett organiseerde een symposium over mentoraat in de opleiding. Er is een opleving van onderzoek te constateren naar de begeleiding van leraren; de Angelsaksische context levert een duidelijke politieke oorzaak voor die opleving. Waar tien jaar geleden voornamelijk procesonderzoek werd gedaan, zien we nu veel meer dat gekeken wordt naar de inhoud van de begeleiding en de aard van de kennis die wordt overgedragen tussen mentor en beginnende leraar. Een sympathieke zoektocht naar manieren om professionele kennis over onderwijzen te registreren, werd ondernomen door Harvard en Mailhos. Als hun op narratieve analyse gebaseerde instrumentarium van de grond komt levert het een belangrijke aanwinst voor onderzoek naar kennis van leraren. Bennetts onderzoek naar begeleiding van leraren is grootschaliger van opzet en heeft tot doel de soorten kennis te identificeren die ontwikkeld worden in de student-mentor relatie. Die kennis draagt bij tot het pedagogisch redeneren van de student-leraar.

Beschrijven van kennis waarover de leraar beschikt, vormt ook het onderwerp van het tweede symposium, over concept-mapping. Dit bevatte vooral methodisch-instrumentele bijdragen. Hoz presenteerde een vrij complexe aanpak waarin kennisdimensies nauwkeurig worden beschreven en anders dan de meeste concept-mapping technieken richt deze aanpak zich niet enkel op frequenties maar ook de inhoud van de cognities. Het instrument maakt

het mogelijk veranderingen in cognities binnen een persoon op te sporen. Zo wordt het mogelijk kennisontwikkeling te volgen. Concept-mapping technieken, zo liet ook Hubers bijdrage zien, zijn weliswaar geschikt om kennis te representeren maar zeggen weinig over kennis, transformatie, kennisherstructurering, en daar gaat het in leren en onderwijzen toch primair om.

Kijkend naar de aard van de bijdragen in thema 4 kan niet gezegd worden dat er nieuwe, prikkelende geluiden te vernemen waren of zelfs dat er aansluiting werd gezocht bij discussies en conceptuele veranderingen in denken over leren en instructie. Er was geen sprake van zelfregulerend leren door studenten in de lerarenopleiding, kennisproductiviteit van lerarenopleiding, sociale geografie van de inductieperiode, de micropolitiek van student-mentor relaties, conceptuele verandering van beliefs en opvattingen, performance assessment en kwaliteit van docenten. Het zijn nieuwe geluiden die niet in thema 4 van de EARLI gehoord werden. Kan het zijn dat in onderzoek naar de leraar en de lerarenopleiding wat over het hoofd wordt gezien?

Constructivisme in reken-wiskundeonderwijs (G. Seegers, Rijksuniversiteit Leiden)

De samenhang tussen leren en instructie als twee met elkaar verbonden elementen is in verschillende onderzoeken op verschillende manieren vormgegeven. Een globale indeling ordent de verschillende toepassingen op een schaal waarbij het ene uiteinde wordt gevormd door docent-gestuurd leren, met aan het andere uiteinde onderwijs waar de eigen inbreng van leerlingen uitgangspunt vormt. In het eerste geval ligt de nadruk op transfer van a priori gedefinieerde vaardigheden, strategieën en concepten. In het andere geval worden de gevolgen voor het leerproces juist beschouwd in termen van een afnemende rol voor (formele) instructie in het leerproces. Ten aanzien van leren in het algemeen, maar ten aanzien van het reken-wiskundeonderwijs in het bijzonder, markeren deze uiteinden tevens posities in de discussie over de meest effectieve aanpak. In deze bijdrage wil ik in het kort nagaan hoe in verschillende bijdragen een positie in deze discussie

wordt ingenomen.

Deze ontwikkeling is gestimuleerd door nieuwe ideeën over leren ('situated learning'), maar vooral door een brede acceptatie van constructivistische benadering van het leerproces. Maar wat betekent acceptatie van het constructivisme? Het begrip constructivisme kent veel interpretaties. We willen daarom dit contrast aanscherpen door gebruik te maken van wat Von Glasersfeld (1987) de twee basisprincipes van het constructivisme noemt: 1. kennis wordt niet passief overgedragen, maar wordt actief verworven; 2. kennis impliceert het structureren van ervaringen, niet het ontdekken van een onderliggende werkelijkheid. Basisidee is dat kennis geen entiteit vormt die kan worden overgedragen, maar dat kennis moet worden geconstrueerd.

Nu is het constructivisme geen onderwijs-theorie. Hoe constructie van kennis in een onderwijskundig kader past moet verder worden uitgewerkt. De beide principes van Von Glasersfeld markeren een epistemologisch en geen didactisch standpunt. Maar acceptatie van het eerste dan wel het tweede principe leidt tot verschillende uitgangspunten bij de didactische uitwerking. Acceptatie van het eerste principe is algemeen. Von Glasersfeld spreekt hier over trivial 'constructivism'. Het tweede principe heeft verdergaande consequenties, en maakt de didactische uitwerking complexer. Het kernpunt is of leren berust op een vorm van transfer van kennis en vaardigheden, dan wel dat het toekennen van betekenis aan ervaringen de kern van het leerproces vormt.

Didactische uitwerkingen van constructivistische principes zijn op een voorspelbare manier verschillend. Acceptatie van het eerste principe laat de assumptie intact dat kennis een a priori gedefinieerde entiteit is die op een of andere manier kan worden overgedragen. Actieve betrokkenheid is voorwaarde voor een effectief leerproces, en instructie is gericht op het vergroten van actieve (cognitieve) betrokkenheid van de leerder. Meer actieve betrokkenheid impliceert minder sturing door de docent, en een grotere verantwoordelijkheid voor de leerder, waardoor het belang van metacognitieve kennis en vaardigheden (in vergelijking met specifieke cognitieve kennis) toeneemt. Dit heeft geleid tot meer gematigde constructivistische benaderingen, waarin nadruk wordt

gelegd op leren in betekenisvolle contexten, en op metacognitieve kennis en vaardigheden als leeropbrengst. In verschillende bijdragen zijn deze meer gematigde benaderingen aan de orde geweest. In het symposium *Activating instruction: Implementation and effects*, gaat Lamberts in op vorm en plaats van instructie. In het symposium *Metacognition and learning to learn: Why train metacognition?* wordt eveneens een aantal fundamentele aspecten aan de orde gesteld. Volet gaat in haar bijdrage in op de rol van instructie bij het trainen van metacognitieve kennis en vaardigheden. Zij benadrukt de noodzaak om metacognitieve vaardigheden te verbinden met specifieke inhouden.

In deze onderzoeken is een vorm van transfer van kennis het uitgangspunt. De verschillende onderzoeken zijn over het algemeen gebaseerd op training en sturing, met nadruk op actieve betrokkenheid van leerders. Daarmee passen ze binnen een gematigde vorm van constructivisme. Acceptatie van het tweede (epistemologisch) principe impliceert dat het criterium voor kennisgroei niet is gelegen in de mate van overeenstemming met een 'objectieve' werkelijkheid (de assumptie die aan het idee van transfer van kennis ten grondslag ligt), maar in de mate waarin kennis functioneel is, en een model oplevert dat 'werkt'. De vraag naar functionaliteit van de kennis betekent een heroriëntatie op aard en kwaliteit van constructies. Leerlingen kunnen bijvoorbeeld kennis ontwikkelen die is gebaseerd op het voornemen zoveel mogelijk antwoorden te geven die de leeraar wil horen. De kennis die de leerling dan ontwikkelt zal bestaan uit opvattingen, principes en regels die zoveel mogelijk goede antwoorden genereren. In de bijdrage van De Corte en Verschaffel aan het symposium *Constraints on knowledge construction and conceptual change: A look across content domains* is het thema aan de orde gesteld dat leerlingen zich vaak een eigen beeld vormen van de bedoeling van het onderwijs en de wijze waarop problemen moeten worden aangepakt. Hier bleek dat leerlingen bij redatiesommen in het algemeen schoolse oplossingen kiezen, zonder over de betekenis van hun oplossing in de context van de opgave na te denken. Hoe didactische principes worden geformuleerd om het ontwikkelen van dit soort 'buggy algorithms' te voorkomen, wordt strikt genomen niet door

de constructivistische theorie voorgeschreven. Een mogelijkheid is om leerlingen te helpen om relevante kennis en procedures zelf vanaf de grond af aan te laten opbouwen. Deze aanpak, die een verwerping inhoudt van het idee dat transfer van kennis leidt tot de gewenste kennis, komt overeen met wat in de realistische theorie van het reken-wiskundeonderwijs het principe van het geleide heruitvinden heet. Met name in de bijdrage van Cobb, Gravemeijer et al. in het symposium *How successful is the 'realistic' approach for mathematics education?* zijn didactische consequenties uitgewerkt. Zij laten zien dat, ook bij acceptatie van het tweede principe van Von Glasersfeld, er 'richtpunten' zijn waarop je het onderwijs kunt richten. Het idee dat acceptatie van het tweede principe de docent met lege handen laat is geen noodzakelijke consequentie.

Literatuur

- Gravemeijer, K. (1995). Het ontwikkelen van 'constructivistisch' rekenwiskundeonderwijs. *Pedagogische Tijdschrift*, 20, 277-292.
- Von Glasersfeld, E. (1987). Constructivism. In T. Husen & N. Postlethwaite (Eds.), *International Encyclopedia of Education* (Suppl. 1). Oxford: Pergamon.

Motivationele, affectieve en sociale processen (J. van der Linden, Universiteit Utrecht)

Het onderzoek op het terrein van leren en instructie kende op het EARLI '95 congres een indeling naar drie domeinen. Het eerste domein betrof *cognition and meta-cognition*. Het tweede ging over *development* en het derde bevatte *motivational, social and affective processes*. In mijn bijdrage aan deze kroniek zal ik uiteenzetten dat deze driedeling een verkeerde indruk zou kunnen geven van het huidige onderzoek naar leerprocessen en de beïnvloeding daarvan. Exemplarisch zal ik enkele bijdragen aan het congres bespreken.

Veel onderwijspsychologen beklemtonen heden ten dage de actieve, constructieve, doelgerichte en cumulatieve aard van menselijk leren en benadrukken tevens dat leren plaatsvindt in een sociale en culturele context. Deze kijk op leren in het algemeen heeft consequenties voor de inrichting van het onderzoek en de thema's die geïdentificeerd worden. Op de eerste

plaats geldt per definitie dat het accent op 'leren in context' de onderzoekers voor de opgave stelt de veronderstelde verwevenheid tussen het cognitieve en het sociale nader te preciseren. De nadruk op context vraagt bovendien een hernieuwde doordenking van de mate van domeinspecificiteit van leren (bijvoorbeeld: transfer binnen en tussen schoolvakken; schools- versus buitenschools leren; contextualisatie versus decontextualisatie). Voorts vormt de veronderstelde cumulatieve aard van leren een uitnodiging om de relatie tussen leren en ontwikkeling in de beschouwingen te betrekken (bijvoorbeeld: de betekenis van voorkennis en van leeftijdsgebonden invloeden) en vraagt de beklemtoning van de doelgerichtheid van leren om onder meer uitbreidingen in de sfeer van perceptie en beleving van de situatie (thematiek: motivatie, emotie en cognitie). Een indeling, waarbij het verstand (domein 1) gescheiden wordt van de wil, het gevoel en sociale invloeden (domein 3) en waarbij aan ontwikkeling (domein 2) een aparte plaats is toegekend, vormt geen goede afspiegeling van hetgeen recent aan onderwijspsychologisch onderzoek plaatsvindt. Een oppervlakkige beschouwer van het EARLI '95 programma op leren en instructie zou dan ook een verkeerde indruk kunnen krijgen van wat op het congres geboden wordt.

Vormen de bijdragen aan het congres een reden om bezorgd te blijven over de vraag of het gepresenteerde onderzoek in voldoende mate aansluit bij de meest recente inzichten over leerprocessen en de beïnvloeding daarvan? Ter beantwoording van deze vraag zal ik beknopt ingaan op enkele presentaties, die deels toebehoorden aan domein 3.

Is er aandacht besteed aan de verwevenheid tussen het cognitieve en het sociale? Ja, zij het dat het sociale op zeer uiteenlopende aggregatieniveaus onderzocht wordt (van microgenetisch in hier en nu situaties tot metacontracten in specifieke culturele omstandigheden) en met eveneens zeer verschillende invullingen van het cognitieve (zoals probleemoplossen, onderhandelen, besluitvorming en planning). Ter illustratie gelden de volgende bijdragen:

M. Scardamalia merkte op dat er slechts sprake kan zijn van een school- en klascultuur

die gericht is op kennisopbouw en begripseren indien geldt dat de lerende een partner is in een intellectueel debat en waarbij de nieuwe technologieën een sociale rol vervullen als uitlegger, tutor of mede-leerder. Engstrom stelde dat voor het verkrijgen van inzicht in de culturele inbedding van het menselijk handelen niet volstaan kan worden met observaties van dit handelen, maar dat bestudeerd moet worden hoe interacties verlopen en welke invloeden daarop inwerken in uiteenlopende culturen bij vergelijkbare interventies om het onderwijs te verbeteren. Pontecorvo wees op de verschillen in argumentatie tussen kind(eren) en opvoeder in schoolse en buitenschoolse situaties en zij vond enige evidentie dat deze verschillen aanleiding zijn tot andersoortige leerervaringen. In een ronde tafel gesprek over factoren in sociale interacties die het leren kunnen bevorderen en die het kunnen belemmeren is als belangrijkste uitdaging genoemd het verbinden van de uiteenlopende theoretische perspectieven (bijv. zich meer richten op de overeenkomsten dan de verschillen tussen Vygotskij en Piaget) en de verschillende methodologieën (bijv. het meer op elkaar zien te betrekken van verschillende niveaus van observatie bij de analyse van eenzelfde interactiefragment en met gebruikmaking van zowel kwantitatieve als kwalitatieve methoden).

Is leren in samenhang met ontwikkeling aan bod gekomen? Voor wat betreft de relatie sociale interactie en leren kwam deze samenhang ter sprake; bij de veronderstelde mechanismen van verandering (socio-cognitief conflict, hulp van een meer gevorderde in aansluiting op resp. Piaget en Vygotskij) en het type taken of situaties (conservatietaken; situaties die naastbij ontwikkeling ontlokken) om ontwikkeling aan af te meten.

Twee symposia kunnen als illustratie dienen. Het ene symposium, *Methodologies for investigating collaborations in learning contexts*, besteedde vooral aandacht aan data-verzamelings- en data-analyse-technieken om het leren in groepsverband in kaart te krijgen. Een opmerkelijke conclusie was dat we er nog niet goed in slagen de interactieprocessen op het spoor te komen die systematisch verbonden lijken met leerwinst. Het andere symposium, *Learning to learn in peer interaction*, besprak

onderzoek naar samenhangen tussen interactionele patronen in onderwijsleersituaties, kenmerkende dialogen daarbinnen en leeruitkomsten.

Is de domeinspecificiteit van leren geproblematiseerd? In ieder geval in de bijdrage van Bereiter, *Beyond situated cognition*, die erop wees dat naast gesitueerd leren ook het niet-gesitueerde leren een kenmerkende menselijke activiteit is en dat elke vorm zijn beperkingen en mogelijkheden heeft om bepaalde doelen te bereiken.

Kreeg de perceptie en beleving van de situatie een plaats? Ja, vooral bij onderzoek naar culturele verschillen en bij studies naar samenhangen tussen cognitie, motivatie en emotie. Zo werd in een symposium, *Cross-cultural issues in academic learning*, onderzoek gepresenteerd dat duidelijk maakte dat er verschillen tussen culturen zijn in opvattingen over leren en hetgeen als een meer of minder passende studieomgeving wordt beschouwd. Het bleek ook dat aanpassingsproblemen van studenten en afstemmingsproblemen tussen docent en student voor een belangrijk deel toegeschreven konden worden aan het in onvoldoende mate rekening houden met dergelijke verschillen. In een structured posterbijeenkomst, *Interactions of cognitive, meta-cognitive and affective factors*, werd in hoofdzaak correlatief onderzoek besproken waarbij een breed scala aan variabelen op cognitief, meta-cognitief en dynamisch-affectief gebied (waaronder variabelen als ervaren moeilijkheidsgraad van de taak, beeld van eigen kunnen) gerelateerd werd aan studieprestaties en studievoortgang. Cognitieve variabelen tonen een sterker verband met schools presteren dan de dynamisch-affectieve en er is veelal sprake van seksespecifieke interacties.

Kortom. Veel bijdragen weerspiegelen de recente onderwijspsychologische aandachtspunten die verbonden geacht kunnen worden aan de opvatting van leren als constructieve activiteit, ingebed in een sociaal-culturele context. De verkeerde indruk die de gekozen driedeling in eerste instantie bij mij opriep is hiermee weggepoetst.

Is hiermee uiteindelijk alles ten goede keerd? Niet helemaal, omdat de nadruk op leren als constructieve activiteit, onderbelicht laat dat er leerprocessen zijn die zich in belang-

rijke mate aan onze bedoelingen onttrekken, die als het ware vanzelf en moeiteloos optreden en die profijtelijk kunnen zijn maar ook nadelig uitvallen (autonome tendenties; incidenteel leren). Ik stel voor naast intentie (of beïnvloeding zoals bedoeld) ook autonomie bij leren (of beïnvloeding zoals bepaald) een plaats te geven op het volgende EARLI congres in 1997 te Athene.

Schrijven en Lezen (J. Andriessen, Universiteit Utrecht)

De 'Special Interest Group' die zich bezighoudt met schrijven, actief gecoördineerd door Rijlaarsdam en Jechle (deze laatste werd opgevolgd door Espéret) was op de conferentie prominent aanwezig. Er waren talloze symposia en presentaties over het onderwerp schrijven, en de groep kwam zelfs voorafgaand aan de officiële opening van de conferentie bijeen (over schrijven in eerste en tweede taal) en ook nog eens op een tijdstip van de dag waarop iedereen al naar huis was (over morfologische aspecten). Het symposium over het schrijven van argumenterende teksten behoorde tot de drukst bezochte van de conferentie. In totaal telde ik acht symposia en postersessies over schrijven.

Volgend jaar wordt in Barcelona (oktober) een conferentie over schrijven gehouden, alsmede een workshop over argumenterende teksten. Er is een boekje uitgekomen waarin namen, adressen en onderzoeksthema's van de leden bij elkaar zijn gebracht (te verkrijgen bij Rijlaarsdam). Het lijkt er op dat het goed gaat met het schrijfonderzoek in Europa.

Schrijfonderzoek kent in elk geval een ontwikkeling in de breedte. De diverse bijdragen waren van sterk wisselende kwaliteit en gaven uiteenlopende uitgangspunten te zien. Leesonderzoek lijkt qua thematiek meer samenhang te vertonen. In dit overzicht wordt eerst het symposium over argumenterende teksten behandeld, omdat dit een aardig beeld geeft van de diverse (maar niet alle) invalshoeken die aan bod kunnen komen bij schrijfonderzoek. Vervolgens wordt een aantal thema's besproken die naar voren kwamen in diverse bijdragen over schrijven en lezen. Ten slotte worden enige omissies signaleerd.

Argumenterende Teksten

In het symposium *Cognitive processes in the learning and teaching of argumentation skills* (georganiseerd door Coirier en Oostdam) waren vijf bijdragen met verschillende invalshoeken. Het vergelijken van dergelijke bijdragen (het doel van het symposium) riep meer vragen op dan antwoorden. Bovendien was de titel van het symposium misleidend: geen van de bijdragen ging echt over onderwijzen en leren.

De symposiumorganisatie had alle deelnemers (inclusief het genodigde publiek) van te voren de papers toegestuurd om zich optimaal op de discussie te kunnen voorbereiden. Bovendien was een bundeltje met commentaren van sommige deelnemers ruim van te voren beschikbaar. De presentaties op de bijeenkomst zelf konden daardoor kort blijven, wat voor degenen in het publiek die zich niet hadden kunnen voorbereiden een nadeel was.

Het onderzoek van Andriessen, Erkens en De Wit ging over de effecten van plaatjes op de coherentie van argumenterende en informerende teksten. Proefpersonen kregen van te voren argumenten voor en tegen een standpunt. Deze informatie was ofwel in de vorm van een illustratie, ofwel in de vorm van een lijstje met argumenten. De effecten van deze twee vormen van informatie op de coherentie van een argumenterende tekst bleken af te hangen van het onderwerp van die tekst. Waar geschreven moest worden over een relatief bekend onderwerp (vakantie in het buitenland of in Nederland), bleek dat de proefpersonen die een lijst met argumenten hadden gekregen betere teksten schreven (in termen van een coherente argumentatiestructuur) dan de proefpersonen die een illustratie hadden gekregen die diezelfde argumenten uitbeelde. In geval van een abstracter onderwerp (natuurgebieden wel of niet beschermen tegen het publiek) bleek juist de illustratie tot een beter resultaat te leiden dan de lijst met argumenten.

De bijdrage van Banks en Apotheloz was in vrijwel alle opzichten anders van karakter. Hier werd uitgegaan van een specifieke theorie, waarin argumentatie als een fundamenteel onderdeel van (de semantiek van) de taal wordt gezien. Wanneer jonge kinderen uitspraken doen als "Je liegt, dus je mag niet meedoen", beschouwen deze onderzoekers het zinsdeel je liegt als een argument voor de daarop volgende

conclusie. Bovendien gaat de theorie er vanuit dat het kind impliciet een rechtvaardiging (topos) voor ogen heeft: leugenaars moeten worden gestraft. De onderzoekers streven ernaar aan te tonen dat dergelijke redeneringen al bij zeer jonge kinderen plaatsvinden. De methode bestond uit het analyseren en interpreteren van natuurlijke dialogen tussen kinderen en tussen kinderen en volwassenen.

Giroud en Schneuwly lieten studenten die Frans als tweede taal moesten leren samenwerken bij het produceren van een ingezonden brief over het krijgen van kinderen door vrouwen op hoge leeftijd. De analyse omvat het zoeken naar relaties tussen uitspraken in de samenwerking dialoog en het gebruik van bepaalde operatoren van modaliteit (mogelijkheid, toestemming, mening, vragen) in de brief. Voor zowel dit als het voorgaande onderzoek gold dat de einduitslag nog niet bekend werd gemaakt.

In de bijdrage van Oostdam en De Gloppe werden de resultaten van vier tests die alle betrekking hadden op identificatie van gezichtspunten en argumenten in complexe zinnen vergeleken volgens een systematisch design. De tests bestonden uit items van enkele zinnen en het was de taak van de proefpersonen (15-18 jaar) voor elk item aan te geven of het een goede argumentatie bevatte en zo ja, om het aanwezige standpunt en/of de argumenten te onderstrepen. Omdat de tests onderdeel uitmaakten van een nationaal peilingsonderzoek had men de beschikking over meer dan duizend proefpersonen. Het bleek dat specifieke indicatoren die een gezichtspunt aangeven de identificatie van de argumentatie bevorderden. Het gezichtspunt bleek verder gemakkelijker te identificeren in items met meervoudige argumentatie. Ten slotte waren items waarin het argument na het gezichtspunt werd gepresenteerd beter te identificeren.

N. Stein kon helaas niet aanwezig zijn op de conferentie. Haar bijdrage was een analyse van een situatie waarin koppels van twee personen met tegengestelde standpunten met elkaar moesten argumenteren, waarbij soms de instructie werd gegeven tot een compromis te komen.

Thema's

Het moge duidelijk zijn dat eenzelfde thema totaál verschillende bijdragen uit kan lokken. De verdere bespreking van activiteiten op de EARLI op het gebied van schrijven en lezen blijkt te doen aan de hand van de volgende thema's: de rol van het geheugen, vergelijkingen tussen landen, en ontwikkeling van lees- en schrijfvaardigheid. Dit overzicht is niet uitputtend, er was nog meer. Aan het eind wordt kort besproken wat opvallend ontbrak.

De rol van het geheugen

Bij het begrijpend lezen is het vormen en bijstellen van mentale representaties onderwerp van een symposium georganiseerd door Van Oostendorp. Een paper van Tapiero en Otero behandelde de manier waarop lezers reageren op inconsistenties in teksten. Van den Broek besprak zijn model voor het begrijpen van teksten. Het model gaat uit van de veronderstelling dat lezers tijdens het lezen trachten zoveel mogelijk de coherentie te bewaren. Gyselink en Tardieu lieten zien dat illustraties kunnen helpen bij het opbouwen van een coherente representatie wanneer deze illustraties de relaties tussen belangrijke elementen in de tekst laten zien. Mullet, Denhiere en Legros onderzochten wat er gebeurt wanneer een reeds gevormd mentaal model moet worden bijgesteld op grond van andere, nieuwe of tegenstrijdige informatie. Het bleek dat, hoewel de groep die andere informatie kreeg de nieuwe informatie trager verwerkte, alle drie de groepen gelijk presteerden op achteraf afgenomen herkennings- en recall-taken. Cole en Goldman lieten aan de hand van hardop nadenkprotocollen zien dat de strategieën voor het opbouwen van een coherente mentale representatie tussen kinderen en volwassenen verschillen. De aandacht van kinderen is meer lokaal, gericht op zinnen of kleine groepjes van zinnen. Volwassenen letten ook op de globale lijn.

In het symposium *Working memory and writing* vergeleek Allal de spellingsprestaties van groep 8 kinderen bij een aantal verschillende taken (schrijf oefeningen, dictee, tekst productie) en bracht deze in verband met een aantal indicaties van het metacognitief niveau. Het bleek onder meer dat de proefpersonen minder fouten maakten tegen werkwoordsverbuigingen bij meer complexe taken. Jeffry en

Underwood onderzochten de invloed van belasting van het werkgeheugen (een reeks cijfers onthouden) op orale en schriftelijke zinsproductie (zinnen afmaken, syntactische priming). Hoe meer getallen moesten worden onthouden, des te minder was de kwaliteit van de ideeëncoördinatie tussen de geproduceerde zinnen. Het bleek dat er geen effect was van geheugenbelasting op syntactische priming, gemeten met een taak waarin de proefpersonen specifieke syntactische formats kregen opgedrongen (He gave X... of He gave X to...).

Rummer en Grabowski bekeken verschillen tussen schrijven, spreken, dicteren en schrijven met onzichtbare inkt, om te kijken wat de rol is van externe informatie, flexibiliteit in het kunnen hanteren van een eigen tempo, de tijd die het duurt om de taal te produceren (schrijven duurt het langst). Proefpersonen keken naar een film en moesten die reproduceren op één van de vier genoemde manieren. De afhankelijke variabelen waren onder meer de hoeveelheid herhalingen, details, de mate van complexiteit en het aantal fouten. Schrijven leidde tot meer details en minder fouten. Geschreven teksten waren meer compleet en bevatten meer details, ook die geschreven waren met onzichtbare inkt.

Fayol besprak een onderzoek waarin een verklaring gezocht werd voor het verschijnsel dat bij jonge kinderen (onder de 10) recall, zinsproductie-span en narratieve recall mondeling beter gaat dan schriftelijk. Komt dit door een te grote geheugenbelasting bij het schrijven of is het een strategische kwestie? De secundaire taken die hier gebruikt werden waren het ritmisch tikken met de vingers en het tekenen van een geometrische figuur.

McCutchen onderzocht revisie-strategieën bij jonge kinderen. De proefpersonen kregen teksten die spelfouten en inhoudelijke fouten tegen de chronologie bevatten. Voor de helft van hen waren de zinnen waarin zich fouten bevonden vet gedrukt. Dit leidde tot een betere correctie van spelfouten maar niet van chronologische fouten. De oorzaak bleek erin te bestaan dat het vet maken van zinnen met fouten leidde tot een vernauwde aandacht op het tekstdeel waarin zich de fout bevond. Dit voorkwam waarschijnlijk de vorming van een adequate macrostructuur.

Internationaal vergelijkend onderzoek

Een interessant Europees project gaat over *Cross-linguistic dimensions in oral and written narrative production*. Hierin wordt de mondelinge en schriftelijke taalontwikkeling van kinderen tussen de 7 en 15 jaar gevolgd voor het Nederlands, Italiaans, Spaans en Duits. Het symposium (organisatie: L. Verhoeven) ging over de ontwikkeling van cohesie in teksten. Pas tegen het einde van de basisschool zijn kinderen in staat geschreven teksten te produceren die het zelfde niveau hebben als hun gesproken verhalen. Er bleken verschillen te zijn tussen de talen in gebruik en functie van cohesieve middelen, bijvoorbeeld voor het handhaven van topische continuïteit. Verschillen in linguïstische achtergrond vormen ook de invalshoek van het onderzoek van Droop en Verhoeven, waarin verschillen in begrijpend lezen tussen basisschoolleerlingen die leren lezen in hun moedertaal (Nederlands) of tweede taal (moedertaal Turks of Marokkaans) worden gemeten. Er werden onder meer effecten gevonden van culturele achtergrond. Tevens bleek dat bij de Nederlandse kinderen de leesprestaties het meest samenhangen met decodeervaardigheid, terwijl bij de andere groepen lexicale en morfosyntactische kennis van het Nederlands het belangrijkste waren.

Ontwikkeling en bevordering van lees- en schrijfvaardigheid

In leesonderzoek is veel aandacht voor zwakke decodeerders. Vergelijkend onderzoek van Bast en Reitsma liet zien dat variatie in toename van leesvaardigheid tussen schoolpopulaties niet te wijten is aan de aanpak van de scholen zelf maar vrijwel volledig terug te voeren is op individuele verschillen. Een aantal bijdragen liet methoden zien om zwakke decodeerders te trainen of te ondersteunen bij het uitvoeren van leestaken. Het onderzoek van Gruwel, Aarnoutse en Van den Bos houdt zich bezig met de rol van auditieve informatie (begrip in luistersituaties) op leesvaardigheid van leerlingen uit het speciaal onderwijs. De eerste resultaten lijken positief. Een veelbelovende methode om zwakke lezers te trainen in decodeervaardigheid maakt gebruik van flitskaarten. Yap paste deze methode toe bij dyslectische kinderen van ongeveer 10 jaar. Deze kinderen moesten klank- en letterclusters her-

kennen in gesproken en geschreven woorden. Het bleek dat verbetering in decodeervaardigheid optrad bij het lezen van pseudoworden die gedurende 200ms werden getoond. Wentink, Van Bon en Schreuder gebruikten de flitskaartmethode om zwakke lezers (8-12 jaar) te trainen in het decoderen van pseudoworden met een of meerdere lettergrepen. Een sterke vooruitgang in benoemsnelheid was het gevolg van de training. Het gebruik van flitskaarten dwingt deze lezers tot snellere verwerking, hetgeen (in elk geval op korte termijn) decodeervaardigheid bevordert. Vergelijking van de rol van syllaben en grafemen gaf bovendien het belang aan van de lettergreep als eenheid van fonologisch decoderen. Van Bon, Tooren en Van Eekelen gaven goede en zwakke lezers een lexicale decisietaak waarin zij uit een lijst woorden op papier de non-woorden moesten doorstrepen. Het bleek dat presentaties op deze taak sterk samenhangen met scores op een begripstaak en in mindere mate met hardop lezen. Bovendien discrimineerde de lexicale decisietaak het beste tussen goede en zwakke lezers. Een dergelijk resultaat leidt tot de overweging dat hardop lezen als indicator van leesvaardigheid, zoals dat nog steeds op school gebeurt, wellicht beter kan worden vervangen door stilleestaken, zoals lexicale decisie.

Leren schrijven is steeds minder een kwestie van veel doen en steeds meer van gerichte instructie en interventie. Het symposium *Effective strategies for teaching writing* toonde een aantal onderzoeken waarin samenwerkend leren centraal stond als methode voor het verbeteren van schrijfvaardigheid. Boscolo en Cissotto rapporteerden positieve effecten bij zeer jonge schrijvers van hardop dicteren bij wat ze zelf schreven. Schriver onderzocht de reacties van jonge schrijvers op teksten van professionele schrijvers. Het doel van dit onderzoek is om schrijvers zich meer bewust te laten worden van de relatie tussen tekst en wat die tekst te weeg kan brengen. Couzijn liet leerlingen medeleerlingen observeren die schrijftaken uitvoerden en die hun teksten beoordeelden. De methode bleek zowel effectiever als efficiënter te zijn ten opzichte van traditioneel onderwijs (= veel schrijven).

Het symposium *Evaluation of text and text pro-*

duction was niet zozeer methodologisch van aard, maar betroffen interventies die beoogden schrijvers zich meer bewust te laten worden van evaluatiecriteria van lezers en beoordelaars. Eigler belichtte in dit kader de resultaten van een meerjarig onderzoek naar het gebruik van externe informatie tijdens het schrijven en de effecten daarvan op de kennis van de schrijver. Boscolo besprak een onderzoek naar de schrijvers eigen subjectieve criteria over de kwaliteit van diens tekst. Lumbelli, Paoletti en Frausin bespraken de effecten van een leergang waarin schrijvers (groepen 6 en 8) lokale coherentiefouten leerden detecteren, door ze te confronteren met protocollen van lezers die moeilijke passages in teksten probeerden te begrijpen.

Lacunes

Wat opvallend ontbrak bij het schrijfonderzoek waren de verbanden met taalkundige benaderingen van het schrijfproces. Voorts was er erg weinig over elektronische tekstverwerking en over software die zich specifiek op schrijven richt. Bij het aanvankelijk lezen ontbrak een benadering die momenteel in de Verenigde Staten opgeld doet, waarbij men probeert zwakke decodeerders via het ervaren van de eigen articulatie meer gevoel voor spellingklank correspondentie bij te brengen. In het algemeen is er nog steeds weinig goed procesonderzoek waaruit gegevens komen die verklaringen kunnen opleveren voor gevonden correlaties of scores op toetsen.

Trends

Schrijfonderzoek presenteerde zich op de EARLI vooral als een beweging die zich naar de buitenkant toe uitbreidt. Wellicht is dit een fase in een veelbelovend leerproces. Een groot aantal bijdragen was beschrijvend van karakter en betrof het volgen van de ontwikkeling van individuele schrijfprocessen of het onderzoek naar samenwerkend schrijven. Opvallend daarbij is de nadruk op de authentieke context van het onderzoek. Vaak treffen we longitudinaal onderzoek aan. In het algemeen is er veel aandacht voor metacognitive factoren. Een tweede trend is die van een toename van experimenteel theorievormend onderzoek waarin onder min of meer gecontroleerde omstandigheden een aantal variabelen gemanipuleerd

wordt. Onderzoekers worden creatiever in het bedenken van situaties waarin zoiets gedaan kan worden. Ten slotte is er een nieuwe ontwikkeling welke betrekking heeft op communicatie door middel van beeldschermen. Hieronder valt het samenwerkend schrijven, maar ook samenwerkend probleemoplossen. Bij het lezen viel het relatief grote aandeel van bijdragen over decoderen op, ten opzichte van de aandacht voor het begrijpend lezen en tekstverwerking. Wel was er een aantal studies dat de aandacht juist richtte op het verband tussen de twee.

Nieuwe informatie- en communicatietechnologieën in onderwijs en onderwijsonderzoek (M. Valcke, Open Universiteit)

EARLI: een eigen benadering van NICT

EARLI-conferenties zijn geen conferenties waarin nieuwe informatie- en communicatietechnologieën (NICT) voorop staan. Ook de 1995-editie van de EARLI-conferentie blijkt hierop geen uitzondering te vormen. Ontwikkelingen, onderzoek of theoretisch werk met betrekking tot NICT worden vooral voorgesteld als tools die nieuwe theoretische inzichten helpen operationaliseren en/of empirische tests van theoretische inzichten ondersteunen. Deze invalshoek is een terechte – expliciete en soms impliciete keuze – die duidelijk een uitvloeisel is van de EARLI-traditie waarin onderzoek over leren en instructie voorop staat. Voorspelbaar – gezien de theoretische oriëntatie van eerdere EARLI-conferenties – was de ruime plaats die hierbij gegeven werd aan constructivistische theorieën, aan problem solving benaderingen en aan adepten van de *situated cognition* referentiekaders. Deze specifieke benadering is herkenbaar in de titels van lezingen, symposia, poster sessies of ronde-tafel meetings. Enkele voorbeelden:

- Powerful learning environments to promote knowledge (re)construction
- Integrated instructional design technology
- The acquisition of skills and concepts: Are there conflicting requirements to computer-based instructional design?
- Interactive graphical representation systems in learning
- Analysing learner-computer interaction:

Lessons from empirical studies

- Learning complex argumentation in a computer-based environment
- The use and effects of (new) media

Verder kon vastgesteld worden dat onderzoek in verband met of op basis van NICT een geïntegreerd onderdeel uitmaakt van meer generieke onderwerpen. Een typisch voorbeeld hiervan was de invited address van D. Leutner (University of Giessen) waarin hij adaptieve leeromgevingen besprak als onderdeel van een lezing over *Adult and Corporate Learning*.

Binnen de EARLI-organisatie bestaat wel een Special Interest Group (SIG) met betrekking tot 'Learning and Instruction with Computers'. Gelukkig volgt ook deze SIG de EARLI-lijn waar het NICT betreft door vooral het geïntegreerde, theoretisch-fundamentele en/of empirisch onderzoek te benadrukken en niet de technologie zelf.

De invloed van de Instructional Design (ID) benadering

Een opmerkelijk en positieve vaststelling is dat heel wat conferentiebijdragen design-principes afleiden of naar voren schuiven die implicaties kunnen hebben voor NICT. Hierbij wordt dus meer en meer een ontwerpbenadering gevolgd die aansluit bij de descriptieve/prescriptieve tradities uit de ID-benadering. Een bijdrage over *geological mapping* is hiervan een sprekend voorbeeld. J. McCracken (Alberta Research Council, Calgary) somde bijv. een reeks design components op die vervolgens werden vertaald in een op NICT gebaseerde leeromgeving. Een ander voorbeeld was de bijdrage van S. Alexander (University of Technology, Sydney). Zij baseerde het design van een interactief multi-media programma op een analyse van misconcepties van studenten over elektriciteitsfenomenen. Het uiteindelijke ontwerp van de leeromgeving laat studenten toe om elektriciteitsbegrippen voor te stellen op het scherm waarbij het systeem de onderliggende conceptuele benadering kon herkennen en eventueel bijsturen. De technology-push blijkt in deze studies weggeëbd te zijn en plaats te hebben ingeruimd voor een duidelijke overwogen inzet van NICT.

Learner control en de invloed van studentkenmerken

Learner control bleek in NICT-gerelateerde bijdragen meer en meer op de voorgrond te komen. Ook het rekening houden met studentkenmerken bij het werken/leren met NICT sluit hierop aan. Deze specifieke benadering kon vastgesteld worden in heel wat richtlijnen, concrete uitwerkingen van NICT of uit de processen/variabelen die in NICT-gerelateerd onderzoek naar voren kwamen:

- pas de interfaces dermate aan dat heel wat keuzen ten aanzien van de vormgeving door de lerende kunnen aangepast worden;
- het oproepen van bepaalde functionaliteiten wordt in handen gegeven van de lerende of de functionaliteiten passen zich aan aan kenmerken van de lerende (adaptive systems);
- de leermaterialen zijn gebaseerd op wat lerenden zelf hebben uitgewerkt; bijv. gebruik schema's, illustraties die eerder door de studenten zijn uitgewerkt; gebruik transcripts van student-reacties als basis voor het uitwerken van feedback;
- het uitleveren van leermaterialen kan aangepast worden aan leerlingkenmerken; bijv. voorbeelden, casussen en opgaven worden gecontextualiseerd in voor die student herkenbare situaties;
- studenten/leerlingen kunnen verschillende leerpaden volgen door het materiaal afhankelijk van hun keuzen en/of studentkenmerken.

Generieke omgevingen

Opvallend blijft dat de voorgestelde NICT-toepassingen steeds 'dedicated' zijn en niet als generieke omgevingen voorgesteld worden die voor een veelheid van domeinen of contexten bruikbaar zijn. In het EARLI-programma kwam in dit verband slechts één voorbeeld naar voren dat toch een algemene ontwikkelomgeving voorstelde om flexibele leermaterialenproductie te realiseren waardoor lerenden, in dit geval studenten hoger onderwijs, cursussen op maat kunnen samenstellen in interactieve leeromgevingen. In deze omgevingen is er bovendien een maximale user-control voorhanden over de gewenste ingebouwde begeleiding bij het studieproces. Een mogelijke verklaring voor de specificiteit van deze NICT-toepassing kan liggen in de aparte onderwijscontext. Het

systeem werd namelijk voorgesteld in de context van open hoger afstandsonderwijs en zelfstudie. Dit onderwijsniveau en vooral het specifieke format ervan kwamen in deze EARLI-conferentie nauwelijks aan bod.

Metaforen

Een aantal conferentiebijdragen onderzocht of toonde exemplarisch wat het belang is van het kiezen van een effectieve/efficiënte metafoor die aan de basis ligt van de interface bij interactieve leeromgevingen. Niegemann (Technische Hochschule, Darmstadt) gebruikte bijv. een bureau-metafoor.

Zuiver constructivistische leeromgevingen: een pas achteruit

In het conferentieprogramma viel op dat minder bijdragen uit de LOGO-traditie stammen. Bij de discussies naar aanleiding van bijv. een bijdrage over Lego-Logo, bleek overduidelijk dat men de minder positieve onderzoeksresultaten en dus het gebrek aan empirische basis met betrekking tot de hypothetische transferwaarde van leren in zuiver constructivistische leeromgevingen als een voldoende basis beschouwt om fundamentele vragen te stellen bij dit soort omgevingen. Uit de discussie blijkt dat men de domeinspecificiteit bij het ontwikkelen van meer algemene probleemoplossingsvaardigheden in de NICT-gebaseerde omgeving wenst vertaald te zien. Het - naïeve? - geloof in de spontane transfer van leerervaringen opgedaan in niet domeinspecifiek-gerelateerde omgevingen is hier opvallend verminderd.

Implicaties van onderzoek voor praktijk

Bij eerdere punten werd gesuggereerd dat de onderbouwing van op NICT gebaseerde leeromgevingen, bijv. vanuit een instructional design benadering, opvallend naar voren kwam. Toch past hier een kritische noot. De specifieke positie van NICT in het EARLI-onderzoek, waarbij vooral toetsing van theoretische punten aan de orde is, roept vragen op naar de praktijkwaarde van heel wat gepresenteerd onderzoek. Bij discussies werd frequent de vraag gesteld naar de praktische relevantie of werd gevraagd naar de 'lessons learned'. In diverse gevallen bleek de onderzoeker hier geen antwoord op te weten, er niet te hebben over nagedacht of werd zelfs expliciet gesteld "I am a re-

searcher". De afstand tussen theorie en praktijk, het toegepast karakter van onderwijsgerelateerd onderzoek is hierdoor in vraag te stellen.

Integratie van communicatieaspecten in de NIT

Zoals reeds eerder werd aangegeven blijken de meeste onderzoekers en deelnemers aan de EARLI-conferentie nauwelijks betrokken te zijn in een onderwijscontext waarin afstandsleren en zelfstudie voorop staan. Dit kan misschien verklaren waarom - in de context van NICT - zo weinig aandacht is besteed aan de communicatiecomponent in de leeromgevingen. Een uitzondering hierop was een symposium over *A European laboratory study of learning from a multi-disciplinary point of view* waarin verschillende bijdragen zich toespitsten op onderzoek m.b.t. *communicative interfaces for collaborative learning* waarin dan ook een scala aan conferencing tools, mailvoorzieningen werden besproken. NICT wordt in de meeste EARLI-onderzoeken gebruikt als een component in een vrij traditionele face-to-face leeromgeving.

Sociale aspecten

Samenhangend met het vorige punt, maar toch enigszins afwijkend kon vastgesteld worden dat 'collaborative activities' bij het gebruik van NICT in leercontexten opvallend veel naar voren kwamen.

Meta-levels of cognitive functioning

In een viertal sessies kon een opvallend parallelisme worden vastgesteld in de opmerkingen die gemaakt werden over de vele verschillende niveaus van cognitief functioneren die we kunnen vaststellen in op NICT-gebaseerde leeromgevingen. In NICT-omgevingen stelt men vast dat studenten soms te sterk cognitief belast worden doordat ze tegelijkertijd aan het domeinhoudelijke vlak, het meta-niveau van hun eigen cognitieve functioneren, de aansturing van de omgeving, het in het oog houden van de resources in de omgeving, etc... kunnen of moeten denken. De onderzoekers stelden hierbij prangende vragen naar het ontwerp van interfaces, de stapsgewijze opbouw van de leerprocessen in de soms zeer complexe NICT-leeromgevingen.

Visual Learning: het potentieel van multi-media overschat

Voorspelbaar bij heel wat op NICT gebaseerde leeromgevingen was het intense gebruik van de nieuwe multi-media mogelijkheden om grafische informatie in de vorm van tekeningen, foto's, schema's tot zelfs videobeelden in de leersituatie te betrekken. Een aantal symposia en afzonderlijke sessies betrokken dit aspect in hun onderbouwing van het design van de leeromgeving. De nieuwe technologische mogelijkheden van de huidige media laten inderdaad toe dat men op een veel visueler manier informatie presenteert en probleemstellingen kan uitwerken. Toch is een scepsis niet misplaatst ten opzichte van de EARLI-presentaties die 'visual learning' in hun vaandel voeren. Veelal werd bij de theoretische onderbouwing gerefereerd naar de verrijking van de 'cognitive representations' die nu mogelijk werden. De vraag kan echter gesteld worden, en ze werd ook op een vrij indringende manier gesteld door G. Salomon (Haifa University), of hier geen misconceptie in het spel is. Hij stelde zelfs dat hier de zeer oude discussie met betrekking tot de psychosemiotics werd 'opgegraven'. Het presenteren van visuele informatie op een computerscherm is duidelijk iets anders en moet duidelijk onderscheiden worden van een interne representatie van kennis door een lende. Rechttoe - rechtaan stellen dat door visuals de mental modellen verrijkt worden is een te naïeve veronderstelling. Bovendien kon ook vastgesteld worden dat de visuele informatie conflicten en/of verstoringen konden veroorzaken in het leerproces of representatiemodel. De twee-dimensionale voorstellingen op het scherm bijv. de interpretatie van klimatologische fenomenen, bleken het leerproces niet steeds te ondersteunen of eenvoudiger te maken

Onderwijseconomische aspecten

De duidelijk onderwijskundig en onderwijspsychologische oriëntatie van de EARLI-conferentie heeft als nadeel dat bepaalde efficiëntie- en effectiviteitsaspecten met betrekking tot NICT niet naar voren komen in de onderzoekopzetten. Gezien de oriëntatie van EARLI als organisatie kan dit geen verrassing zijn. Toch stellen deelnemers aan sessies telkens weer vragen naar bijv. het onderwijseco-

nomische rendement of de input-output verhoudingen. Indien we dit vertalen in meer algemene termen zou dit kunnen pleiten voor het instellen van een nieuwe Special Interest Group (SIG) of het opnemen van een nieuw conferentiethema in het programma van de volgende bijeenkomst in Athene (1997).

NICT als tool voor EARLI-conferenties

We kunnen ons ook afvragen hoe NICT een rol hebben gespeeld bij het opzetten en uitvoeren van de EARLI-conferentie. Nuttig was bijv. dat voor de eerste keer alle abstracts van conferentiebijdragen op schijf werden geplaatst. Tijdens de conferentie waren heel wat NICT-tools voorhanden in een aparte ruimte, maar het gebruik ervan was miniem. De presentaties zelf waren nauwelijks multi-mediaal. Er kan zelfs gesteld worden dat het bordgebruik en het gebruik van een overheadprojector vrij zwak was. Onderzoekers (i.v.m. leren en instructie) blijken bepaalde vaardigheden van hun onderzoeksobjecten zelf niet steeds te beheersen. In vergelijking met andere conferenties werd nog geen parallelle virtuele EARLI-conferentie via e-mail of discussion lists opgezet. Een idee voor de volgende conferentie? Eventueel als een opwarmer enkele weken voor de eigenlijke 'face-to-face' conferentie in Athene? Interessant was de aankondiging van EARLINet wat het elektronisch forum voor EARLI zou moeten worden. Intekenen op de elektronische - gestructureerde - discussielijst kan via listserv@nic.surfnet.nl (of listserv@hearn.bitnet). Meer informatie kan men krijgen bij de beheerder van de lijst: dochy@OUH.NL. Eerstdaags komt er ook een EARLI-pagina op WWW.

NICT en EARLI: voor wie interessant?

Het hier geschetste kritische perspectief op de recente EARLI-conferentie roept meteen de vraag op voor wie de conferentie interessant is geweest? Voor wie vanuit een praktijkperspectief NICT-toepassingen zocht was en is EARLI duidelijk niet het juiste forum. IFIP, SIGCAL, CAL, ICTE, CALISCE, etc. zijn organisaties/conferenties die duidelijk meer de technologie en haar toepassingsmogelijkheden voorop plaatsen. Ook bij deze conferenties komt onderzoek naar voren maar zijn de conferentiedoelstellingen niet primair op onderzoeksrapportering en verdere theorievorming gericht.

Wie zoekt naar een discours waarin de onderbouwing of de design benadering van NICT als leeromgeving aan bod komt, is binnen EARLI wel op zijn plaats en kan een groeiend aantal discussiepartners vinden.

Aan deze kroniek werkten mee: *J. Andriessen, J. van der Linden, J.M. Pieters, G. Seegers, H.H. Tillema, M. Valcke, M.V.S. Veenman.*

De eindredactie werd verzorgd door
J.M. Pieters

Boekbespreking

H. van Keulen

Making Sense: Simulation-of-Research in Organic Chemistry Education

Academisch proefschrift Universiteit Utrecht

CD-B Press, Utrecht 1995

213 pagina's f 30,- ISBN 90-73346-24-X

Het Utrechtse Centrum voor Didactiek van Wetkunde en Natuurwetenschappen verricht onderzoek naar onderwijs- en leerprocessen binnen de β -vakken, waarbij het accent wordt gelegd op de vakinhoud. De onderzoeksvragen hebben veelal betrekking op onderwerpen die voorkomen in de programma's in de bovenbouw van het HAVO en VWO, of in de onderbouw van universitaire opleidingen.

Het proefschrift van Van Keulen is gebaseerd op een chemiedidactisch onderzoek. Hierin stond een practicum voor eerstejaars scheikundestudenten centraal. Onderzocht is in hoeverre de studenten door het volgen van dit practicum worden voorbereid op het verrichten van zelfstandig wetenschappelijk onderzoek. Van Keulen is niet de eerste die de Utrechtse scheikundepractica tot onderwerp van zijn promotieonderzoek heeft gemaakt. Hij werd voorafgegaan door De Jager (1985), Goedhart (1990) en Elzenga (1991). De reden om deze practica zo grondig te onderzoeken, is gelegen in een structureel probleem dat zich wereldwijd voordoet bij (universitaire) chemische practica. Doorgaans zijn deze practica gericht op het leren onderzoeken of experimenteren. In werkelijkheid leert de student vooral om proeven uit te voeren door een (gedetailleerd) voorschrift te volgen. In de literatuur wordt dan ook gesproken van een 'cookbook'-probleem. Studenten leren dus wel om technieken te hanteren, maar niet om zelfstandig en met gebruikmaking van theoretische inzichten een (chemische) onderzoeksvraag aan te pakken. En dat laatste is uiteraard wel de bedoeling van een universitaire opleiding tot chemicus.

Het kookboekprobleem is niet eenvoudig op te lossen. Van Keulens voorgangers hebben pogingen ondernomen om de betreffende practica anders in te richten, maar stuiten daarbij op problemen van verschillende aard. Practicumopdrachten met een 'open' karakter bleken voor

studenten veelal onhanteerbaar, voor docenten onbegeleidend en voor de practicumorganisatie (veiligheid!) onaanvaardbaar (De Jager, 1985). Ook bleken de meeste studenten niet in staat om theorie, waarover zij reeds met succes een tentamen hadden afgelegd, aan te wenden bij het ontwerpen en uitvoeren van een experiment (Elzenga, 1991). Ten slotte werden experimenten die gericht waren op het leren onderzoeken, gefrustreerd door practicumdocenten die oplossingen aandroegen waar studenten zelf naar hadden moeten zoeken (Goedhart, 1990).

Van Keulen treedt niet klakkeloos in de voetsporen van zijn voorgangers. Voordat hij zich concreet stort op het op microniveau onderzoeken van een gedeeltelijk heringericht eerstejaars practicum, kadert hij zijn onderzoek helder in. In hoofdstuk 2 geeft hij een genuanceerde analyse van het bestaande curriculum. Hij bespreekt daarbij ook de verdiensten van 'voorschriftproeven', maar betoogt ten slotte dat een curriculum waarin kennis en technieken afzonderlijk aangeboden worden, uitgaat van de naïeve veronderstelling dat de voor het verrichten van onderzoek noodzakelijke integratie van kennis en vaardigheden 'vanzelf' optreedt.

In hoofdstuk 3 laat hij zien dat het bestaande curriculum (impliciet) gebaseerd is op een objectivistische filosofie, hetgeen tot uiting komt in een hoge mate van overdracht van kennis en technieken. Van Keulen stelt terecht dat deze filosofie te kort schiet: studenten leren wel om antwoorden ('feiten') te geven, maar niet om vragen te stellen. In plaats van zich aan te sluiten bij het constructivisme, dat al enige tijd erg populair is in kringen van onderzoekers van β -onderwijs, zoekt van Keulen zijn heil in de hermeneutiek. Deze fenomenologische stroming is gebaseerd op de stelling dat wetenschappelijke kennis wordt gevormd door ervaringen en interpretaties van onderzoekers, die hierover consensus trachten te bereiken. Taal, context en tijd spelen hierbij een belangrijke rol. Een hermeneutische benadering is in natuurwetenschappen niet erg gangbaar, maar Van Keulen acht deze juist geschikt als het gaat om het leren en 'begrijpen' van natuurwetenschappen.

Op basis hiervan zet Van Keulen de stap naar de methodologie van zijn onderzoek (hoofdstuk 4). Nadat hij heeft vastgesteld dat veel traditionele methoden van onderwijsonderzoek geba-

seerd zijn op het objectivisme, kiest hij voor een cyclische onderzoeksmethode waarin de termen 'ervaren', 'interpreteren', 'begrijpen' en 'toepassen' centraal staan. Dit betekent: Van Keulen ontwerpt onderwijssituaties die hij observeert. Op basis van de interpretatie van de verkregen gegevens moet bij hem als onderzoeker begrip ontstaan voor het handelen en spreken van de studenten, hetgeen hij kan toepassen in een nieuw onderwijsontwerp, waarna een nieuwe cyclus begint. Dit proces gaat door totdat 'verzadiging' is opgetreden, dat wil zeggen dat de onderzoeker geen nieuwe inzichten meer verwerft.

Dezelfde hermeneutische principes vormen ook het uitgangspunt voor de onderwijsontwerpen (hoofdstuk 5 t/m 7). Van Keulen heeft zich gericht op enkele synthesproeven in het eerstejaars practicum. Hierbij maken studenten een bepaalde hoeveelheid van een opgegeven chemische verbinding. Hij heeft een alternatieve werkvorm ontwikkeld die hij 'simulatie-van-onderzoek' noemt: studenten doen ervaringen op bij het chemisch handelen, interpreteren deze in het licht van theoretische noties en trachten in onderlinge besprekingen tot consensus en daarmee tot begrip te komen. Het accent ligt dus op het zelfstandig argumenteren en experimenteren en daarmee worden de essentiële onderzoeksactiviteiten 'gesimuleerd'.

In het kader van een bespreking in dit tijdschrift acht ik het niet zinvol om verder uit te weiden over de chemische thema's (reactietypen en soorten verbindingen) die zijn onderzocht. Van Keulen trekt zelf de conclusie dat hij na enkele onderzoekscycli in staat is om op kwalitatieve wijze te voorspellen hoe het onderwijsproces zal verlopen. Ten aanzien van het 'begrijpen' door studenten rapporteert hij wisselend succes: de interactie tussen chemisch handelen en argumenteren verloopt productief zolang de theoretische noties die met een bepaalde synthese verbonden zijn nog betrekkelijk eenvoudig zijn (ruwweg: VWO-eindexamen niveau). Zo gauw de theorie complexer wordt (eerstejaarscollege organische chemie) 'vervalen' studenten tot het maken van de gevraagde verbindingen zonder hierbij theoretische noties te betrekken.

slaagd om onderwijs te ontwerpen dat studenten daadwerkelijk aanzet tot zelfstandig experimenteren en argumenteren. De werkvorm 'simulatie-van-onderzoek' biedt in mijn ogen perspectief voor het vervangen van meerdere kookboekproeven. Toch wil ik wel enkele kanttekeningen bij het onderzoek plaatsen.

Ten eerste de vraag voor welke studenten de conclusies geldig zijn. In zijn concluderende paragrafen schrijft Van Keulen over 'the students'. Toch heeft hij de leerprocessen van slechts enkele studenten, voornamelijk via protocolanalyses (op basis van audio-opnames van groepsbesprekingen) onderzocht. De keuze voor het verzadigingsprincipe, gericht op het begrip van de onderzoeker, staat in mijn ogen op gespannen voet met de wens om algemene uitspraken te doen over studenten.

Ten tweede de docentenproblematiek, waar Van Keulen zelf ook op wijst. Docenten zijn op het onderzochte practicum dikwijls beginners qua doceerervaring (veelal AIO's) en bouwen ook weinig ervaring op. Van Keulen heeft dit ervaren als een ernstige belemmering: docenten konden op deze wijze geen volwaardige hermeneutische cycli doorlopen. Van Keulen heeft de docenten weliswaar geïnstrueerd via een handleiding, maar bekritiseert deze aanpak zelf vanwege het overdrachtskarakter. Zijn voorkeur voor de hermeneutiek lijkt in dit geval een excuus om de docent grotendeels buiten beschouwing te laten. Ik vind dit niet terecht omdat Van Keulen van te voren op de hoogte was van de docentenproblematiek, zowel specifiek voor wat betreft het door hem onderzochte practicum (zie onderzoek Goedhart), alswel algemeen op basis van onderzoek naar de rol van docenten bij onderwijsinnovaties. Zijn suggestie dat het management van het practicum en de faculteit een ander beleid zouden moeten voeren ten aanzien van de bemensing van de practica, is nogal gemakzuchtig.

Deze kritiek laat onverlet dat Van Keulen een zeer relevant onderzoek heeft verricht, dat concreet inspeelt op de actuele discussie over academische vaardigheden in het universitaire onderwijs. Bovendien hanteert hij een prettig leesbare schrijfstijl (in het Engels!), die maakt dat het grootste deel van dit proefschrift ook voor niet-chemici toegankelijk is. Een sterk punt is ook de fundering van het onderzoek in wetenschaps-

Is het kookboekprobleem nu eindelijk opgelost? Dat nog niet, maar ik meen dat de oplossing wel dichterbij is gekomen. Van Keulen is er in ge-

theoretische en methodologische zin. Ofschoon de keuze voor de hermeneutiek bekritiseerd kan worden (de vergelijking met het constructivisme (p. 62) bijvoorbeeld is nogal zwart/wit en het verschil met andere cycli, zoals de experiëntiële leeracyclus van Kolb, is in mijn ogen slechts graduëel), wordt overtuigend aangetoond dat een onderwijsontwerp op basis van ervaring, interpretatie, begrip en toepassing geschikt kan zijn voor onderwijs dat gericht is op het verwerven van academische vaardigheden. Onderzoek naar de toepassing van een dergelijke aanpak binnen andere disciplines kan zeer waardevol zijn.

J.H. van Driel

Literatuur

- Elzenga, H.E. (1991). *Kwaliteit van kwantiteit*. Utrecht: CD-β Press.
- Goedhart, M.J. (1990). *Metten: Normen en waarden*. Utrecht: Rijksuniversiteit Utrecht.
- Jager, H. de (1985). *Leren synthesiseren – symbiose van ambacht en wetenschap*. Utrecht: Rijksuniversiteit Utrecht.

Mededelingen

Velon-Congres 1996

Het jaarlijkse congres van de Vereniging Lerarenopleiding Nederland (Velon) vindt plaats op 18 en 19 maart 1996 te Leeuwarden. De organisatoren zijn de Noordelijke Hogeschool Leeuwarden en de Christelijke Hogeschool Noord-Nederland.

Voor nadere informatie: Mevrouw Tj. Idsinga, Noordelijke Hogeschool Leeuwarden, tel.: 058 293 42 05.

Studiedag Classificatie en diagnostiek in de jeugdzorg

Deze studiedag die georganiseerd wordt door de PAO Gedragswetenschappen in samenwerking met het Nederlands Instituut voor Zorg en Welzijn (NIZW) en de Nederlandse Werkgroep Classificatie en Registratie, wordt op 22 maart 1996 te Ede gehouden.

Voor inlichtingen: PAO, Postbus 325, 2300 AH Leiden, tel.: 071 5143 143.

Congres Onderwijs, opvoeding en sociale integratie

Dit congres dat georganiseerd wordt door het SCO-Kohnstamm Instituut van de U.v.A. zal plaatsvinden te Amsterdam, op 17 en 18 oktober 1996.

Voor nadere informatie: Drs. H. Lington, SCO-Kohnstamm Instituut, Wibautstraat 4, 1091 GM, tel.: 020 5251 000, fax: 020 5251 000.

Promoties

Rijksuniversiteit Leiden
Promovendus: Drs. F. J. G. Meijerink
Titel proefschrift: A non-linear structural relations model
Promotor: Prof. dr. L. J. Th. van der Kamp
Datum: 18 januari 1996

Universiteit van Amsterdam
Promovendus: Drs. N. P. J. Kousemaker
Titel proefschrift: Zoeken, vinden, zorgen delen.
De ontwikkeling van een praktijkparadigma

voor onderkenning en pedagogische preventie van psychosociale problematiek in de Jeugdhulpverlening

Promotor: Prof. dr. A. T. G. van Gennep
Datum: 2 februari 1996

Katholieke Universiteit Nijmegen
Promovendus: Drs. H. H. A. Verschuur
Titel proefschrift: Ethische grondslagen van de morele opvoeding, een utilistisch perspectief
Promotores: Prof. dr. A. W. van Haaften en Prof. dr. H. J. L. van Luijk
Datum: 14 februari 1996

Rijksuniversiteit Groningen
Promovendus: Drs. J. M. Doornenbal
Titel proefschrift: Ouderschap als onderneming; moeders en vaders over opvoeden in de jaren negentig
Promotores: Prof. dr. C. Brinkgreve en Prof. dr. K. B. Koster
Datum: 15 februari 1996

Rijksuniversiteit Leiden
Promovendus: Drs. I. Bakkenes
Titel proefschrift: Professional isolation of primary school teachers; a task-specific education
Promotor: Prof. dr. F. K. Kievit
Co-promotor: Dr. J. Imants
Datum: 5 maart 1996
Tijd: 14.15 uur

Universiteit van Amsterdam
Promovendus: Drs. K. J. Polder
Titel proefschrift: Implementatie van wetgeving in onderwijs
Promotor: Prof. dr. A. M. L. van Wieringen
Datum: 19 maart 1996
Tijd: 15.00 uur

Universiteit van Amsterdam
Promovendus: Drs. M. Overmaat
Titel proefschrift: Schrijven en lezen met tekstschema's. Effectief onderwijs in schriftelijke taalvaardigheid in de bovenbouw van het voortgezet onderwijs
Promotores: Prof. dr. C. M. de Gloppe en Prof. dr. B. H. A. M. van Hout-Wolters
Co-promotor: Dr. G. C. W. Rijlaarsdam
Datum: 28 maart 1996
Tijd: 15.00 uur

Rijksuniversiteit Leiden

Promovendus: Drs. J. M. A. Börger

Titel proefschrift: Transfer enhancement and knowledge acquisition in teacher education

Promotor: Prof. dr. F. K. Kieviet

Co-promotor: Dr. H. H. Tillema

Datum: 25 april 1996

Tijd: 14.15 uur

Inhoud andere tijdschriften

Comenius

15e jaargang, nr. 4, 1995

Vijf onderzoeksbenaderingen: een inleiding, door J. Jongkind

Het historisch-sociologisch onderzoek van het

Nederlands onderwijsbestel, door J. Wolthuis

De klas, de klankmethode en het ontstaan van de

achterblijver, door I. Weijers

Muziekonderwijs, waarom eigenlijk? door C.

Koopman

Sensitiviteit als pedagogische categorie: een po-

ging tot verheldering, door M. de Wolff en R.

van der Veer

Emancipatie in de fabriek en op het kantoor,

door E. Offers en A. Jorna

Pedagogisch Tijdschrift

20e jaargang, nr. 4/5, 1995

Themanummer Curriculum, constructivisme en

authentiek leren

Inleiding. In de leer bij de constructivisten, door

J. Carpay en J. Terwel

Constructivisme en curriculum, door K. Th.

Boersma

Cultuuroverdracht als reconstruerende activiteit,

door B. van Oers

Het ontwikkelen van 'constructivistisch' reken-

wiskundeonderwijs, door K. Gravemeijer

Authentiek leren in de basisvorming, door H.

Franssen, E. Roelofs en J. Terwel

Hoe nieuw is het constructivisme? Een nabe-

schouwing vanuit een onderwijspedagogische

invalshoek, door J. Carpay

Pedagogisch Tijdschrift

20e jaargang, nr. 6, 1995

Pedagogisch commentaar, door N. Deen

Reikwijdte van transfer van een korte training in

Inductief Redeneren, door W. Tomic

Traditie en vernieuwing in het lees- en taalonderwijs voor jonge meertalige kinderen, door E. Wagenaar

Besprekingsartikel Kritische onderwijssociologie in Noord-Amerika - 15 jaar na dato, door G.

J. J. Biesta

Tijdschrift voor Onderwijsresearch

20e jaargang, nr. 4, 1995

Beoordeling van het motorische niveau en de schrijfmotoriek van basisschoolleerlingen door leerkrachten, door B. C. M. Smits-Engelsman, G.

P. van Galen en C. G. J. Michels

Tijdsbesteding in combinatieklassen en enkelvoudige klassen. Een descriptief onderzoek naar

besteding van de beschikbare tijd op basisscho-

len, door M. Kral, M. Voeten en W. Nederlof

Is strategische hulp effectief in een aardrijks-

kunde-computerles? door T. van der Zijpp

Leerstijlenonderzoek gerepliceerd; de samen-

hang tussen leerstijlen en intelligentie, door V. V.

Busato, F. J. Prins, C. Hamaker en K. H. Visser

Het relatieve belang van sociaal milieu en etni-

sche herkomst voor de verklaring van onderwijs-

achterstanden, door G. W. J. M. Driessen

Gedragstoetsen voor professionele gespreks-

vaardigheden: Ontwikkeling, betrouwbaarheid

en validiteit, door H. T. van der Molen en W.

Vermeulen

Tijdschrift voor Onderwijsresearch

21e jaargang, nr. 1, 1996

Themanummer: Het jonge kind in onderzoek

Inleiding: Onderwijs aan jonge kinderen nog

nauwelijks onderzocht, door E. Harskamp en G.

Reezigt

Onderzoek naar vormgeving en effecten van het

onderwijs aan jonge kinderen in Nederland, door

G. Reezigt

Stimulering van geletterdheid bij kleuters, door

J. van Kuyk en L. Verhoeven

Ontwikkeling van tweetaligheid bij allochtone

kleuters, door L. Verhoeven en G. Narain

Actuele opvattingen over onderwijs aan jonge

kinderen. Een onderzoek onder leerkrachten van

groep 2 in het basisonderwijs, door L. M. Ste-

vens en E. G. Harskamp

Onderwijs aan jonge kinderen: historische ont-

wikkelingen, door G. W. Meijnen

en verder:

Notities en Commentaren

Rekeninhoudelijke kritiek op CEB-repliek, door A. Treffers

Treffers schiet mis, door W.J. van der Linden en M. A. Zwarts

Maar raakt wel zijn eigen voet, door W.J. van der Linden en M. A. Zwarts

Tijdschrift voor Orthopedagogiek

34e jaargang, nr. 11, 1995

Stimuleren van motivatie en zelfvertrouwen. Het effect van een programma voor de mytyschool, door E. S. Kunnen

Effect woordlengte op herkenning woorden en non-woorden bij volwassen dyslectici, door M. I. Geerlings, W. Kraay, A. Bouma en J. W. van Strien
Gedragsbeoordeling van kinderen met enco-
prepsis, door J. D. Bosch, W. A. van Hoogstraten
en L. T. ten Brink

Tijdschrift voor Orthopedagogiek

34e jaargang, nr. 12, 1995

Jongeren in justitiële behandelingsrichtingen, door L. Boendermaker en C. Verwers

Ambulante hulp aan jeugdigen, door J. D. van der Ploeg en E. M. Scholte

Dyslexie. Afbakening en behandeling. Advies van een commissie van de Gezondheidsraad, door A. J. J. M. Ruijsenaars en D. C. M. Gersons-Wolfensberger

Tijdschrift voor Orthopedagogiek

35e jaargang, nr. 1, 1996

In memoriam Dr. Jan E. Kappers, 1946-1995, door D. J. Bakker e.a.

Behandeling verstandelijk zwakbegaafde zedendeliquenten, door F. Bruinsma

Kinderen met een verstandelijke belemmering, in het bijzonder Down's syndroom, in het reguliere onderwijs: motieven van ouders, door G. W. de Graaf

Gezinsbeleid: Nationale redding of nationale kwaal? door J. Gerris

Ontvangen boeken

Barritt, L. S., *An elementary school in Holland. Experiment in educational practice*. International Books, Utrecht, 1996, f 35,-.

Braet, A., *Schrijfvaardigheid Nederlands. Een praktische didactiek voor de bovenbouw van HAVO en VWO*. Uitgeverij Coutinho, Bussum, 1995, f 24,50.

Branger, J. D. C., *Twee eeuwen voor de klas. Lerarenopleiding basisonderwijs Haarlem*. Uitgeverij De Vrieseborch, Haarlem, 1995, f 28,50.

Brentjens-Horsten, D. M. M., *De dynamiek van verandering en stabiliteit: Een historische vergelijkende gevalstudie van ontwikkelingen in het basisonderwijs tussen 1969 en 1990* (dissertatie). Academisch Boeken Centrum, De Lier, 1996.

Bruininks, A. C., *Ze doen het niet expres. Een structuurverlenende aanpak bij aandachtsstoornissen en hyperactiviteit*. Uitgeverij SWP, Utrecht, 1995, f 32,50.

Druine, N., Haak, E., Lagerweij, N., Wielemans, W., Dewulf, L., *Verwantschap en verscheidenheid. Het secundair onderwijs in Vlaanderen. Het voortgezet onderwijs in Nederland. Een vergelijking*. Nederlandse Taalunie, 's-Gravenhage, 1995.

Grol, M. van, *Kinderopvang Plus. Een handleiding voor opvoedingsondersteuning in kindercentra*. Uitgeverij SWP, Utrecht, 1995, f 24,50.

Heuvel-Panhuizen, M. van den, *Assessment and realistic mathematics education* (dissertatie). CD-β Press, Center for Science and Mathematics Education, Utrecht, 1996.

Hox, J. J., Meulen, B. F. van der, Janssens, J. M. A. M., Laak, J. J. F. ter, Tavecchio, L. W. C., *Advances in family research*. Thesis Publishers, Amsterdam, 1995, f 45,-.

Kooten, A. H. van, & Petegem, J. J. G. van, *Interne begeleiding. Zorgactiviteiten planmatig coördineren binnen het schoolteam*. Uitgeverij Intro, Baarn, 1995, f 29,90.

Laak, J. J. F. ter, *Psychologische diagnostiek. Inhoudelijke en methodologische grondslagen*. Swets & Zeitlinger, Lisse, 1995, f 59,-.

Pijnenburg, H. M., Rijswijk, C. M. van, & Veerman, J. W. (Red.), *Pedagogisch Jaarboek 1995*. Uitgeverij Eburon, Delft, 1995, f 39,50.

Wal, J. van der, Mooij, I. de & Wilde, J. de, *Identiteitsontwikkeling en leerlingbegeleiding*. Uitgeverij Coutinho, Bussum, 1996, f 47,50.

Wotherspoon, T., & Jungbluth, P. (Eds.), *Multicultural education in a changing global economy: Canada and The Netherlands*. Waxman, Münster/ New York, 1995, DM 49,90.

Inhoudsopgave

BIBLIOTHEEK DE
RIJKSUNIVERSITEIT
UTRECHT

Pedagogische Studiën Jaargang 73, nummer 2/1996

R. Appel en A. Vermeer

Uitbreiding van de Nederlandse woordenschat van allochtone leerlingen in het basisonderwijs 82

R. H. Hofman

Effecten van het verschil in bestuursvorm voor het openbaar en bijzonder basisonderwijs 93

F. J. Prins, V. V. Busato, C. Hamaker en K. H. Visser

Een bijdrage tot de validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen 108

W. A. M. Vollebergh

Autoritarisme en etnocentrisme bij leerlingen uit lagere opleidingsniveaus: instemmingstendens of opinie? 123

Discussie

C. J. W. Meijer

Realisering Weer Samen Naar School-beleid afhankelijk van wijziging bekostigingssysteem 136

A. J. J. M. Ruijsenaars

Weer Samen Naar School: als het doel op is, beginnen we aan de middelen 140

G. H. van Gemert

Prikkels tot zorgzaam onderwijs? 143

C. J. W. Meijer

Reactie op de bijdragen van A. J. J. M. Ruijsenaars en G. H. van Gemert 146

Kroniek

J. M. Streumer (Red.)

Annual Meeting American Educational Research Association
18-22 april 1995, San Francisco 149

Boekbesprekingen 158

Mededelingen 165

Uitbreiding van de Nederlandse woordenschat van allochtone leerlingen in het basisonderwijs

R. Appel en A. Vermeer

Samenvatting

De achterstand van allochtone leerlingen in het basisonderwijs en hun doorstroming naar lagere vormen van vervolgonderwijs is voor een groot deel te wijten aan hun geringere woordenschat in het Nederlands. Om de woordenschat van allochtone kinderen te vergroten is een onderwijsprogramma voor groep 1 tot en met 4 samengesteld op basis van beschikbare leermiddelen. In dit artikel worden de resultaten van een evaluatie-onderzoek gepresenteerd dat aan de invoering van dit programma was gekoppeld. In een longitudinaal onderwijs-experiment hebben we in twee opeenvolgende groepen nagegaan of de woordenschat in het Nederlands van allochtone leerlingen ($n=57$) substantieel vooruitgang in het experimentele programma door hun scores op curriculum-afhankelijke en -onafhankelijke toetsen te vergelijken met die van leerlingen van controlegroepen, zowel van andere als van dezelfde scholen ($n=53$ resp. 39). Tevens is nagegaan of het woordenschatprogramma een breder effect had op hun begrip- en leesvaardigheid in groep 5. Ten slotte is de omvang van hun woordenschat vergeleken met die van Nederlandse leeftijdsgenoten. De resultaten maken duidelijk dat het mogelijk is de woordenschat in het Nederlands van allochtone leerlingen een enorme impuls te geven. Eind groep 4 bereiken de kinderen die het experimentele programma gevolgd hebben weliswaar niet het niveau van hun Nederlandse leeftijdsgenoten, maar lopen zij in woordenschat toch één à twee jaar voor op kinderen uit dezelfde etnische groep die het programma niet gevolgd hebben.

Van Kemenade genoemd) onder meer het volgende.

Al sinds het begin van de zeventiger jaren wordt door de Nederlandse overheid beleid gevoerd gericht op opheffing van achterstanden in het onderwijs voor specifieke groepen leerlingen. Daaraan is vanaf 1979 toegevoegd een aantal maatregelen om met name de educatieve achterstanden van leerlingen uit culturele minderheden te bestrijden. (...) Het is bijzonder teleurstellend dat evaluatie-onderzoek laat zien, dat de geïnvesteerde gelden en de gespendeerde energie nog onvoldoende effect opbrengen in de zin van sterk verbeterde schoolprestaties van allochtone leerlingen.

De onderzoeksgegevens waaraan in dit citaat wordt gerefereerd, zijn met name afkomstig uit het landelijk OVB-onderzoek. Uit de rapportage van Tesser, Mulder en Van der Werf (1991) over dit onderzoek blijkt dat allochtone leerlingen zowel voor taal als voor rekenen veel lagere scores behalen dan autochtone leerlingen uit lager sociaal milieu, die op hun beurt weer (iets) lager scoren dan autochtone leerlingen uit hoger sociaal milieu. Een ander opmerkelijk en alarmerend gegeven is dat die achterstand van autochtone kinderen uit laag sociaal milieu na het zesde leerjaar gelijk blijft, maar dat de achterstand van allochtone leerlingen ten opzichte van hun Nederlandse klasgenootjes zelfs nog toeneemt. Hoewel uit het vervolgonderzoek naar voren komt dat de achterstand van allochtone leerlingen constant blijft, wordt daarbij meteen opgemerkt dat het in deze longitudinale vergelijking wel gaat "om de 'betere' 1.9-leerlingen. Vergeleken met de twee autochtone groepen is een relatief groot deel van de leerlingen die in groep 4 is getoetst, blijven zitten of naar het speciaal onderwijs verwezen. Wanneer we het over de mate van onderwijsachterstand hebben, moet dat eigenlijk worden verdisconteerd." (Mulder, 1994, p. 84). Mulder stelt verderop (Mulder, 1994, p. 109)

Inleiding

Op de eerste pagina van haar rapport *Ceders in de tuin* schrijft de Commissie Allochtone Leerlingen in het Onderwijs (ook wel de Commissie

dat het gegeven dat de allochtone leerlingen op OVB-scholen het in 1992 iets beter doen dan in 1988 "voor het grootste deel verklaard kan worden door de langere verblijfsduur in Nederland".

Waar dit alles in resulteert is duidelijk. Allochtone leerlingen stromen over het algemeen door naar lagere vormen van voortgezet onderwijs, dat zij bovendien vaker dan autochtone leerlingen zonder diploma verlaten, waarna ze vervolgens op de arbeidsmarkt relatief weinig kansen hebben (Tesser, 1993).

Nederlandse taalvaardigheid is een belangrijke factor als het gaat om het leveren van schoolprestaties. Een goede taalbeheersing in het Nederlands is uiteraard niet de enige, maar wel een noodzakelijke voorwaarde voor schoolsucces. Het is zonder meer duidelijk dat allochtone leerlingen op dit punt hun schoolcarrière beginnen met een flinke achterstand ten opzichte van hun autochtone klasgenootjes, omdat ze over het algemeen thuis opgroeien met de taal van de eigen etnische groep. Boogaard, Damhuis, De Gloppe en Van den Bergh (1990) deden een peilingsonderzoek naar de Nederlandse taalvaardigheid van Surinaamse, Antilliaanse, Marokkaanse, Turkse en Nederlandse leerlingen aan het eind van de kleuterperiode. Uit de taalvaardigheidsmeting met een aantal deeltoetsen van de Taaltoets Allochtone Kinderen (TAK) (Verhoeven & Vermeer, 1986) kwam een redelijk consequent patroon naar voren: Surinaamse en Antilliaanse kinderen presteren iets meer dan één standaarddeviatie onder het gemiddelde van de Nederlandse groep, de Marokkaanse kinderen anderhalve standaarddeviatie en de Turkse kinderen twee standaarddeviaties. De prestaties op een aantal spreektaken laten dezelfde rangorde zien tussen de onderscheiden taalgroepen. Gemiddeld over de spreektaken presteren de Surinaamse en Antilliaanse leerlingen iets meer dan een halve standaarddeviatie lager dan het gemiddelde van de Nederlandse groep, de Marokkaanse leerlingen ruim één standaarddeviatie en de Turkse leerlingen bijna anderhalve standaarddeviatie.

Ander onderzoek naar de ontwikkeling van Nederlandse taalvaardigheid van allochtone leerlingen versterkt het sombere beeld dat al uit de studie van Boogaard c.s. naar voren komt. Verhoeven en Vermeer (1989), die scores op de

TAK cross-sectioneel onderzochten bij kinderen van 4 tot 9 jaar oud, laten zien dat er sprake is van een blijvende taalachterstand. Die geldt vooral de Nederlandse woordenschat van allochtone kinderen. In absolute zin, dat wil zeggen in termen van het aantal beheerste woorden, neemt die achterstand zelfs toe in de loop van de basisschool. Verhoeven en Vermeer (1992) rapporteren over vervolgonderzoek met de Taaltoets Allochtone Kinderen voor de Bovenbouw. Daarin wordt de bovengenoemde trend nog eens bevestigd. Verhoeven en Vermeer extrapoleren de scores op hun toetsen naar het totaal beheerste woorden. Dan blijkt voor de receptieve of passieve woordenschat het volgende te gelden. Turkse en Marokkaanse kinderen van vier jaar en zes maanden kennen gemiddeld 1000 woorden receptief en hun Nederlandse leeftijdsgenootjes (uit laag sociaal milieu) gemiddeld ruim 3000 woorden, dus een verschil van ruim 2000 woorden. Turkse en Marokkaanse leerlingen van twaalf jaar en zes maanden hebben een receptieve woordenschat van ruim 10.000 woorden en Nederlandse leerlingen uit die leeftijdsgroep hebben een receptieve woordenschat van 15.000 woorden.

Relatief gezien neemt de achterstand dus af, maar absoluut gezien (in aantallen woorden, van 2.000 naar 5.000) toe. Daarbij ligt het voor de hand om te veronderstellen dat de 5000 woorden die Nederlandse leerlingen meer kennen vooral 'moeilijke woorden' zijn, betrekkelijk laag frequente woorden die allochtone kinderen niet makkelijk in het alledaags taalgebruik zullen oppikken, maar die op school wel belangrijk zijn, bijvoorbeeld bij het lezen van teksten voor de zaakvakken. Met betrekking tot het laatste moet ook het volgende schrijvende gegeven nog worden gesignaleerd: op 8-jarige leeftijd, als ze bijvoorbeeld teksten voor wereldoriëntatie moeten gaan lezen, is de receptieve woordenschat van allochtone leerlingen ongeveer net zo groot als die van autochtone kleuters bij hun intree in de basisschool.

Al deze onderzoeksresultaten hebben betrekking op de *omvang* van de woordenschat, waarbij bestudeerd is of kinderen de betekenis van een woord (meestal de hoofd- of kernbetekenis) receptief of produktief kennen. Er is echter ook onderzoek uitgevoerd waarin het accent lag op de kwaliteit van de betekenistoekenning.

Zo wordt uit de studie van Verhallen (1994) duidelijk dat Turkse 9-jarige en 11-jarige kinderen aan eenvoudige Nederlandse woorden minder betekenisaspecten toe kennen dan hun Nederlandse leeftijdsgenootjes. Bovendien blijkt de betekenis-toekenning van Turkse kinderen in vergelijking met Nederlandse kinderen minder paradigmatisch en meer syntagmatisch van aard. Dat wil zeggen, een woord als *neus* wordt door de Turkse kinderen eerder in verband gebracht met *oor* dan met *lichaams-deel*. De paradigmatische relaties, die wijzen op een verder en hiërarchisch ontwikkeld lexicaal systeem, komen vaker voor bij Nederlandse kinderen. Onderzoek van Karumma (1992) heeft laten zien dat Marokkaanse en Turkse leerlingen in groep 7 en 8 van de basisschool minder betekenissen van polyseme woorden zoals *vleugel* ('de vleugel van een vogel', 'de vleugel van een gebouw') kennen dan hun Nederlandse klasgenootjes. De twee genoemde studies wijzen erop dat allochtone leerlingen niet alleen in *kwantitatieve*, maar ook in *kwalitatieve* zin een flinke achterstand hebben in hun Nederlandse woordenschat.

Waarom hebben we nu in het bovenstaande zo'n sterk accent gelegd op *woordenschat*? Daar zijn de volgende redenen voor:

1. De achterstand in woordenschat is, in vergelijking met andere linguïstische niveaus, het meest prominent. Dat is vermoedelijk een gevolg van het feit dat het bij die andere linguïstische niveaus (klanken, woordvorming, zinsbouw) altijd om een betrekkelijk beperkt aantal regels gaat, terwijl de woordenschat bijzonder omvangrijk is.

2. Uit verschillende studies is gebleken dat er een sterke samenhang bestaat tussen enerzijds de lexicale vaardigheid en anderzijds andere taalvaardigheden van tweede-taalverwerwers (zie o.a. Kerkhoff, 1988; Verhoeven & Vermeer, 1989). De correlatie zegt uiteraard niets over een eventueel causaal verband, maar het lijkt ons aannemelijk om te veronderstellen dat er bijvoorbeeld eerder een lijn loopt van lexicale vaardigheid naar grammaticale vaardigheid dan andersom: wie veel woorden kent in een taal, zal die taal eerder en makkelijker gebruiken en op die manier zijn grammaticale vaardigheid verder kunnen ontwikkelen.

3. Verschillende studies constateren een samenhang tussen lexicale vaardigheid en onder-

wijsresultaten. Anderson en Freebody (1981) geven een samenvattend overzicht van onderzoek naar de relatie tussen lexicale en andere vaardigheden. De correlaties tussen de scores op een aantal vocabulairetests en scores op een aantal verschillende IQ- en schoolvorderingentests blijken te lopen van .71 tot .98. Anderson en Freebody gaan nog iets dieper in op de relatie tussen lexicale vaardigheid en begrijpend lezen, en stellen op basis van een groot aantal studies dat woordkennis in sterke mate samenhangt met begrijpend lezen (zie ook Aarnoutse & Van Leeuwe, 1988). Ook hier is weer niet duidelijk wat de richting van het eventuele causale verband is, maar ons inziens is de veronderstelling gewettigd dat die eerder loopt van woordenschat naar begrijpend lezen dan omgekeerd. Mogelijk is er ook sprake van een wisselwerking in die zin dat er een flinke woordenschat nodig is voor begrijpend lezen, en dat middels (begrijpend) lezen de woordenschat vervolgens verder kan worden ontwikkeld.

Bovenstaande overwegingen leiden tot de conclusie dat het van groot belang is om zo vroeg mogelijk in het basisonderwijs te werken aan uitbreiding van de woordenschat van allochtone leerlingen. Een bijkomend argument is gelegen in de volgende constatering: wie veel woorden kent, kan makkelijker de betekenis van onbekende woorden afleiden, c.q. nieuwe woorden leren. Een leerling die in een zin van tien woorden twee onbekende woorden tegenkomt, zal waarschijnlijk de zin niet begrijpen, terwijl een leerling die slechts één woord niet kent, vermoedelijk die zin op z'n minst voor een deel zal begrijpen en ook nog ongeveer de betekenis van dat ene onbekende woord zal kunnen afleiden. Die betekenisafleiding is dan de basis voor de uitbreiding van de woordenschat met dat betreffende woord. Anders gezegd: meer woorden kennen, betekent meer woorden leren.

Om de woordenschat van allochtone kinderen uit te breiden is er een speciaal programma nodig omdat de plaats van woordenschatuitbreiding in regulier onderwijsmateriaal veel te gering is. Wij hebben zo'n programma voor groep 1 tot en met 4 samengesteld op basis van beschikbare materialen. Aan de invoering van het programma is het evaluatie-onderzoek gekoppeld waarvan we

hier samenvattend de resultaten presenteren. Hieronder zullen we het programma nader beschrijven; aan het eind van die paragraaf komen de onderzoeksvragen aan de orde.

1 Het onderwijsprogramma en de onderzoeksvragen

Bij de samenstelling van het onderwijsprogramma zijn we uitgegaan van de volgende drie principes voor adequaat woordenschatonderwijs voor (jonge) kinderen:

- woorden moeten worden aangeboden in een betekenisvolle context;
- de aangeboden woorden moeten voldoende worden herhaald, omdat anders de betekenisovername niet genoeg wordt verankerd;
- als ze nieuw worden aangeboden, moeten woorden zo veel mogelijk worden ondersteund met visueel materiaal, vooral voor jonge kinderen.

Zoals gezegd, zijn we bij de samenstelling van het programma uitgegaan van bestaand materiaal dat werd geselecteerd op basis van de bovenstaande drie principes, met als bijkomend uitgangspunt dat er voldoende woorden moesten worden aangeboden. De doelstelling van het programma was namelijk dat allochtone leerlingen per jaar *ruim duizend woorden* zouden leren, zodat ze aan het eind van de vierjarige periode wat betreft hun woordenschat op hetzelfde niveau zouden zitten als Nederlands-talige kinderen van dezelfde leeftijd. Anders gezegd: het programma moest in principe voldoende woorden bevatten met het oog op de voorgenomen inhaalmanoeuvre.

Het programma was samengesteld uit bestaande methoden Nederlands als tweede taal (NT2) aangevuld met prentenboeken. Van die prentenboeken moest de tekst ook op bandjes staan, zodat de kinderen die zelfstandig herhaald zouden kunnen beluisteren. De aanvulling met prentenboeken bleek vooral noodzakelijk omdat veel NT2-materiaal veel te weinig woorden bevatte om de doelstelling van ruim duizend woorden per jaar te realiseren. Per leerjaar zag het programma er als volgt uit.

1. *Knoop het in je oren* (NT2) plus zestien prentenboeken;
2. *Laat wat van je horen* (NT2) plus twintig prentenboeken;

3. *Van Horen en Zeggen*, deel 2 en *Mondeling Nederlands bij Veilig Leren Lezen* (beide NT2) plus tien prentenboeken;
4. *Allemaal Taal A* (NT2) en acht prentenboeken.

We willen met nadruk stellen dat het niet onze bedoeling was om het *ideale programma* op te stellen. Daarvoor was het nodig geweest om eerst te onderzoeken welke vier à vijfduizend woorden in de vier jaar geleerd zouden moeten worden en vervolgens zelf leermateriaal te ontwikkelen (of te laten ontwikkelen) op basis van die lijst met doelstellingswoorden. Een dergelijke opzet lag ver buiten onze mogelijkheden. Onze ambitie was om na te gaan of de Nederlandse woordenschat van allochtone leerlingen in de eerste vier groepen van het basisonderwijs substantieel kan worden uitgebreid. Het programma dat we ten behoeve daarvan samenstelden, was op dat moment vanuit ons perspectief het best mogelijke, alhoewel we op sommige onderdelen ervan ook wel kritiek hadden en hebben. Zo blinkt de NT2-methode *Laat wat van je horen* niet uit wat betreft duidelijkheid en overzichtelijkheid van illustraties en zijn niet altijd alle woorden in de prentenboeken even relevant, bijvoorbeeld voor groep 1 *brommen* in de betekenis van 'brommend zeggen', en *buizerd*.

Op basis van het bovenstaande hebben we de volgende onderzoeksvragen geformuleerd.

1. Gaat de Nederlandse woordenschat van allochtone leerlingen die gedurende de eerste vier jaar van het basisonderwijs het experimentele programma volgen meer vooruit dan de woordenschat van allochtone leerlingen die het programma niet volgen?

2. Heeft het woordenschatprogramma een breder effect, bijvoorbeeld op begrip- en leesvaardigheid?

3. Is het mogelijk om allochtone leerlingen een zodanig onderwijsaanbod (in de vorm van het experimentele programma) te geven dat ze ruim 1000 woorden per jaar leren en halen ze dan het niveau van hun Nederlandse leeftijdsgenootjes?

2 Opzet van het onderzoek

In een longitudinaal onderwijsexperiment hebben we nagegaan of met het invoeren van bovengenoemd programma in groep 1 tot en met 4 van het basisonderwijs de drie gestelde onderzoeksvragen bevestigend beantwoord kunnen worden. Het experiment is eind 1989 op een zestal scholen gestart in groep 1. Naast leerlingen op die experimentenscholen volgden we vier jaar lang ook leerlingen op zeven controlescholen. Bovendien is op de experimentenscholen een tweede cohort gevolgd van leerlingen die het jaar daarna, in 1990, in groep 1 gekomen waren. Ten slotte zijn de prestaties van de leerlingen op de experimentenscholen aan het eind van de vier jaar vergeleken met die van de leerlingen die twee groepen hoger zaten op dezelfde school, en die niet het experimentele onderwijsprogramma gevolgd hadden (het zgn. cohort 0), en wel op het moment dat ook zij vier jaar onderwijs genoten hadden. Zo is het mogelijk de eerste vraag te beantwoorden door de leerlingen van de experimentenscholen te vergelijken met zowel leerlingen van andere als van dezelfde scholen. Om de tweede vraag te beantwoorden is aan alle leerlingen, toen ze in groep 5 zaten, in de maand maart een taak voor begrijpend lezen voorgelegd. Het antwoord op de derde vraag wordt verkregen door de onderzoeks-kinderen gestandaardiseerde, curriculum-onafhankelijke woordenschattoetsen voor te leggen die een indicatie kunnen geven van de omvang van hun woordenschat en op basis waarvan ze vergeleken kunnen worden met een normgroep van autochtone leeftijdsgenoten.

2.1 Scholen en proefpersonen

Van de zes experimentenscholen staan drie scholen in de regio Tilburg, één school in Zwijndrecht en twee scholen in Amsterdam. Het relatieve aandeel allochtone leerlingen loopt van tien procent (Oisterwijk) tot honderd procent (Zwijndrecht). De twee scholen in Tilburg (veertig en zestig procent allochtonen) en de twee in Amsterdam (negentig procent allochtonen) zitten daar tussenin. In het eerste jaar waren er op deze scholen 14 groepen bij het experiment betrokken. Als onderzoekskinderen beschouwen we alleen die kinderen die thuis een andere taal spreken dan het Nederlands en

die bij het eerste meetmoment (november 1989) nog geen vijf jaar waren, en die voor het eerst in groep 1 zaten. De onderzoekskinderen zijn daarmee bijna uitsluitend van Turkse en Marokkaanse afkomst. De prestaties van deze kinderen van de experimentenscholen zijn vergeleken met die van de kinderen uit in totaal zeven controlescholen, drie uit Amsterdam en vier in Tilburg. Een deel van de controlescholen had zich aanvankelijk aangemeld als experimentenschool, maar was afgefallen bijvoorbeeld omdat ze ons programma te sterk voorschrijvend vonden. Afhankelijk van de betreffende school, werd een of meer keren per jaar overleg gevoerd met de school. Verder is jaarlijks via een schriftelijke enquête aan de leerkrachten gevraagd in welke mate ze het programma uitvoerden, hoeveel tijd de verschillende onderdelen kostten, hoe de deelname van de kinderen was, welke beoordeling ze gaven aan de gebruikte leermiddelen, et cetera (zie voor een verslag bijvoorbeeld Appel & Vermeer, 1990b).

De experimentenscholen hebben in meer of mindere mate het voorgeschreven programma uitgevoerd, waarbij er grote verschillen zijn in de door de scholen geïnvesteerde tijd, namelijk gemiddeld ongeveer tussen twee uur en vier uur per week. Eén van de experimentenscholen week vanaf het tweede jaar zeer sterk van de andere af: de leerkrachten bleken na groep 1 nauwelijks nog het voorgeschreven programma te volgen. Net zoals de controlescholen besteedde deze school wel extra aandacht aan NT2-leerlingen, deels met dezelfde materialen als in het experiment, maar ze kon niet meer als experimentenschool beschouwd worden. Deze school is verder in de analyses meegenomen als controleschool.

Na ruim vier jaar onderwijs voor zowel cohort 1 als cohort 2 konden we 110 leerlingen die aan alle criteria voldeden, in de analyse betrekken: 57 leerlingen van de experimentenscholen (32 van cohort 1, 25 van cohort 2), en 53 leerlingen van de controlescholen. Deze 110 hadden alle curriculum-onafhankelijke toetsen gemaakt, en negen van hen hadden hooguit een van de acht curriculum-afhankelijke toetsmomenten gemist. Bij het eerste meetmoment was de gemiddelde leeftijd van beide groepen 4 jaar en 7 maanden, en waren de seksen in beide groepen gelijkelijk verdeeld. De controlegroep

op dezelfde experimentescholen ('cohort 0') bestond uit 39 leerlingen. Dit cohort 0 was bij afname van de laatste toets in groep 4 gemiddeld een half jaar ouder dan cohort 1 bij de afname van de laatste toets in groep 4. Alle kinderen waren onderinstromers en kwamen uit de laagste sociaal-economische klassen (beroepsniveau 1 en 2).

2.2 Toetsen

In het onderzoek is zowel gebruik gemaakt van curriculum-afhankelijke als van curriculum-onafhankelijke toetsen. De eerste hebben we uiteraard zelf samengesteld, en voor de tweede zijn de woordenschattoetsen uit de Taaltoets Allochtone Kinderen (TAK) (Verhoeven & Vermeer, 1986) gebruikt. Door de vooruitgang in woordkennis van de experimentgroep te vergelijken met die van de controlegroep is het mogelijk de effecten van ons programma te meten, zowel wat betreft de in het programma aangeboden woorden (jaarlijks) als woordkennis in het algemeen (over de gehele periode). Om begrip- en leesvaardigheid te meten is een half jaar na het experiment, toen de leerlingen in groep 5 zaten, in maart een voorversie van Tekstleestaak 1 uit de TAK-bovenbouw (Verhoeven & Vermeer, 1993) afgenomen. In totaal hebben de leerlingen dus negen meetmomenten (mm) gehad: aan het begin en eind van elk schooljaar (mm1 en mm2 in groep 1, tot en met mm7 en mm8 in groep 4) en in het voorjaar van groep 5 (mm9) de Tekstleestaak. Met uitzondering van de Tekstleestaak, waarmee schriftelijk-klassikaal werd getoetst, zijn alle toetsen mondeling-individueel in 20 à 30 minuten afgenomen.

Van de curriculum-afhankelijke toetsen is aan het begin en eind van elk schooljaar een voormeting en een nameting gehouden. Ze meten de kennis van de in dat jaar in het programma aangeboden woorden: in groep 1 de *Prog1*-test, in groep 2 de *Prog2*, etc. De programmatoetsen zijn deels als receptieve taak geconstrueerd volgens het bekende multiple-choice-principe: de proefleider zegt een woord, het kind krijgt vier plaatjes te zien en moet het correcte plaatje aanwijzen, en deels productieve taken waarbij de leerling een plaatje getoond wordt en hij of zij een door de proefleider aangeboden zin moet aanvullen ('Dit is een ...', 'De jongen is aan het ...', etc.). In de eerste drie

groepen bevatten deze toetsen steeds 100 items; in groep 4 bestond deze uit 50 items. In deze groep werd bovendien een definitietetaak (*Prog4b*) gegeven met items als: 'Wat is *sjoken*?', 'Wat is een *kozijn*?', etc. De betrouwbaarheden van deze curriculum-afhankelijke toetsen waren goed: Cronbachs alpha was voor *Prog1*, *Prog2*, *Prog3* en *Prog4a* en *Prog4b* (definitietoets) bij de voormetingen respectievelijk .92, .93, .88, .81 en .88.

Van de curriculum-onafhankelijke toetsen is de Passieve-woordenschattoets van de TAK onder andere afgenomen op mm1 (begin groep 1) en mm8 (eind groep 4), en de Actieve-woordenschattoets op mm3 (begin groep 2) en eveneens op mm8 (eind groep 4). Door middel van extrapolatie van de scores op deze toetsen kan tevens een indicatie gegeven worden van de receptieve en productieve woordvoorraad van de kinderen. Zie voor een beschrijving van deze toetsen en verantwoording van de extrapolatie verder Verhoeven en Vermeer (1992). Op basis hiervan was het mogelijk de derde vraag te beantwoorden, namelijk of door middel van het programma de leerlingen ruim 1000 woorden per jaar zouden kunnen leren en of ze zo het niveau van hun autochtone leeftijdsgenootjes kunnen halen.

2.3 Procedure

In het experiment is uitgegaan van implementatie van bestaande onderwijsleermiddelen (methoden en prentenboeken), begeleid met longitudinaal onderzoek ('aptitude treatment interaction design'). Met behulp van variantie-analyses van de scores op de curriculum-afhankelijke en -onafhankelijke toetsen is vervolgens geprobeerd een antwoord te geven op de gestelde vragen.

Met het oog op de eerste vraag is nagegaan of er bij elk van de vier curriculum-afhankelijke *Prog*-toetsen sprake is van een jaarlijkse significante vooruitgang voor de leerlingen van experimentescholen in vergelijking met die van de controlescholen. Tevens is nagegaan of dat ook geldt voor de curriculum-onafhankelijke toetsen, door de eerste en de laatste afname van beide toetsen te vergelijken. Hierbij zijn SPSSX-MANOVA-analyses gedraaid op de verschillen tussen voor- en natoetsen. Hoewel voorheen in sociaal-wetenschappelijk onderzoek covariantie-analyses als betrouw-

baarder gezien werden, zijn er inmiddels goede argumenten aangedragen voor het gebruik van 'rechte' verschijscores (vgl. Willet, 1987). Met het oog op de tweede vraag is (in SPSSX-ANOVA) nagegaan of er sprake is van significante verschillen in scores op de begrippeleleestaak tussen experiment- en controlescholen. Wat betreft de derde vraag worden de geëxtrapoleerde scores op de woordenschattoetsen van de TAK eind groep 4 vergeleken met die uit het normeringsonderzoek van de TAK, met andere woorden, of de omvang van de woordenschat van de leerlingen van de experimentgroep na vier jaar het niveau bereikt heeft van hun Nederlandse leeftijdsgenoten. Bij de beantwoording van de tweede en de derde vraag worden de leerlingen van de experimentescholen niet alleen vergeleken met die van de controlescholen, maar ook met die van cohort 0, de leerlingen op dezelfde experimentescholen die het programma niet gevolgd hebben.

3 Resultaten

In Tabel 1 wordt een overzicht gegeven van de scores van de experiment- en controlescholen op de curriculum-afhankelijke toetsen. De scores van de twee deelttoetsen in groep 4 zijn opgeteld. De maximale score op elke toets is 100.

Zoals af te lezen is uit Tabel 1, gaan beide groepen op elk van de vier programmatoetsen vooruit. Daarbij scoort de experimentele groep zowel op de voor- als natoets steeds significant hoger. Behalve op mm3, is ook de variantie in scores steeds significant, met andere woorden, de experimentele groep is heterogener dan de controlegroep. Toch is er geen toename van he-

terogeniteit over de tijd heen, en gaan de experimentescholen op de programmatoetsen meer vooruit dan de controlescholen. In vier afzonderlijke variantie-analyses met Prog1, Prog2, Prog3 en Prog4 als afhankelijke variabele waren de factoren 'tijd' en 'exp/contr' significant, bij Prog1, Prog3 en Prog4 ook de interactie ('tijd' $\times$ 'exp/contr', sig. of F respectievelijk $p = .015$; $p = .000$; $p = .001$). Beide groepen gaan kortom vooruit in de loop van de tijd, en in groep 1, 3 en 4 gaan de experimentescholen significant meer vooruit dan de controlescholen.

Toch is de toename van de scores lager dan we verwachtten. Voor een deel is dat het gevolg van het feit dat niet alle experimentescholen het hele jaarprogramma afgewerkt hadden, voor een ander deel bleken de toetsen relatief moeilijk te zijn. Bij verschillende items van de Progtoetsen konden we constateren dat de kinderen een woord wel deels verworven hadden, maar niet zodanig dat dit tot uitdrukking kwam in de score. Zo werd voor het item 'tulp' bij de voortoets op mm1 elk willekeurig plaatje aangewezen (ook de palmboom), omdat de kinderen zich bij dat woord niets konden voorstellen. Maar bij de natoets op mm2 wisten de meeste kinderen dat een tulp een bloem is, en werd de palmboom niet meer aangewezen, maar wel naast de tulp ook andere bloemen die afgebeeld waren in de afleiders. De meeste kinderen hadden dus wel iets verworven, maar aan de scores was dit niet af te lezen.

Anderzijds is de 'leerwinst', d.i. het percentage woorden dat de leerlingen geleerd hebben van het aantal bij de eerste meting onbekende woorden, niet erg afwijkend van wat in andere onderzoeken doorgaans gevonden wordt, namelijk gemiddeld zo'n 25 à 35% (zie Appel & Vermeer, 1990a, p. 45). Op Prog1, Prog2, Prog3 en Prog4 was de leerwinst bij de experimentgroep resp. 33%, 29%, 27% en 21%; bij de

Tabel 1

Gemiddelden en standaarddeviaties van experiment- en controlescholen op de voor- en nametingen van de programmatoetsen (maxima steeds 100)

		Prog1		Prog2		Prog3		Prog4	
		mm1	mm2	mm3	mm4	mm5	mm6	mm7	mm8
experiment	gem	46.0	64.3	42.8	58.3	30.0	49.2	29.4	44.5
	sd	16.4	15.5	18.9	16.7	14.4	14.8	13.4	18.4
controle	gem	34.6	48.0	31.3	46.5	20.2	31.2	18.7	28.2
	sd	110.4	14.1	15.4	14.0	10.6	10.9	6.7	11.5

Tabel 2

Gemiddelden en standaarddeviaties van de experiment- en controlescholen op de Passieve- en Actieve-woordenschattoon van de TAK (max. = 60 resp. 98), en op de Tekstleestaak (max. = 40), voor alle cohorten

		Passief Woord		Actief Woord		Tekstleestaak
		mm1	mm8	mm3	mm8	mm9
experiment (n=57)	gem	12.1	67.3	12.5	31.5	25.0
	sd	10.7	13.3	8.6	9.8	5.4
controle (n=53)	gem	6.2	52.7	7.2	22.8	19.8
	sd	5.9	13.4	6.1	6.6	6.0
cohort 0 (n=39)	gem				24.7	19.2
	sd				8.3	4.7

controlescholen was die 16%, 22%, 14% en 11%. Wanneer een gedeeltelijke verwerving van een woord in de scores verwerkt wordt (bijvoorbeeld: het kind heeft wel geleerd dat een tulp een bloem is, maar weet niet precies wat voor soort bloem), dan loopt de leerwinst in de experimentgroep gemiddeld op tot boven de 50%. Elders (zie Appel & Vermeer, 1994) hebben we geprobeerd verklaringen te vinden voor de verschillen tussen 'goed' en 'slecht' geleerde woorden. Deze werden gevonden in bijvoorbeeld factoren als de frequentie van voorkomen van de woorden in de teksten, het aantal verschillende boeken waarin ze voorkomen en de mate van ondersteuning door de (verbale) context dan wel het aantal illustraties waarop het woord is afgebeeld.

Naast deze programmatoetsen zijn de leerlingen de Actieve- en Passieve-woordenschattonen van de TAK afgenomen, alsmede een voorversie van de Tekstleestaak van de TAK-bovenbouw (Verhoeven & Vermeer, 1993) in groep 5. In Tabel 2 staan de resultaten van de eerste en de laatste meting van de experiment- en controlescholen, en de scores op de Tekstleestaak.

Ook op de curriculum-onafhankelijke toetsen gaan de experimentescholen meer vooruit dan de controlescholen, zoals af te leiden valt uit Tabel 2. In een variantie-analyse (SPSSX-MANOVA) met Actieve Woordenschat als afhankelijke variabele waren de factoren 'tijd' en 'exp/contr' significant, als ook de interactie tussen deze factoren ($F=9.63$; $df=1,108$; $p=.003$). Ook bij de Passieve Woordenschat was dit het geval ($F=15.38$; $df=1,108$; $p=.000$). Beide groepen gaan kortom vooruit in de loop van de tijd, maar de experimentescholen gaan significant meer vooruit dan de controlescho-

len. Dit betekent dat het onderwijsprogramma een positief effect heeft op de totale woordenschat van de kinderen, en dat niet alleen de onderwezen woorden geleerd worden. Daarbij wordt de standaarddeviatie niet veel hoger bij een toenemend gemiddelde, hetgeen erop duidt dat alle kinderen profiteren van het programma, en er geen achterblijvers zijn binnen de experimentgroepen. Bovendien blijkt uit de scores van het cohort 0 op de Actieve-Woordenschattoon dat er geen sprake is van een 'school-effect'. Omdat zowel op deze curriculum-onafhankelijke als op de curriculum-afhankelijke toetsen de experimentele groep op de voormetingen significant hoger scoort dan de controlegroep (zie Tabel 1 en 2), zou men kunnen concluderen dat hier sprake is van een vergelijking tussen 'goede' en 'slechte' scholen. De leerlingen van de controlegroep binnen de experimentescholen scoren echter op de Actieve-woordenschattoon nauwelijks hoger, en op de Tekstleestaak zelfs lager dan die van de andere controlescholen, ondanks het feit dat ze gemiddeld een half jaar ouder zijn. De eerste onderzoeksvraag kan met andere woorden positief beantwoord worden: de leerlingen die het experimentele programma gevolgd hebben, gaan in woordenschat meer vooruit.

Zoals te verwachten, zijn er tussen scholen grote verschillen in scores en de mate van vooruitgang. De 'leerwinst' op sommige experimentescholen ligt rond de 50%, op andere rond de 25%. Daarbij scoren de Amsterdamse scholen steeds het laagst, zowel bij de experiment- als de controlescholen. Toch is het niet van belang of een school 'zwart' (meer dan 50% allochtoon) dan wel 'wit' (minder dan 50% allochtoon) is: de school die bijvoorbeeld op de Actieve-Woordenschattoon het hardst vooruit-

gaat (nl. van 3.8 naar 33.2) heeft 100% allochtonen, de school die het hoogst scoort in groep 4 op deze toets (39.5) heeft 60% allochtonen. Ook uit eerdere variantie-analyses bleek dat er geen significante interactie was tussen de hoofdeffecten 'wit/zwart' $\times$ 'tijd' (zie Appel & Vermeer, 1990a, p. 43).

Uit de scores op de Tekstleestaak (zie Tabel 2) komt naar voren dat het experimentele programma een breder effect heeft dan alleen op woordenschat. Ook hier scoren de experimentescholen hoger dan de controlescholen en de controlegroep binnen de school. Uit een variantie-analyse (SPSSX-ANOVA) met de Tekstleestaak als afhankelijke variabele bleek het verschil tussen experiment- en controlescholen significant ($F=13.997$; $df=1,69$; $p=.000$), evenals dat tussen de leerlingen van dezelfde school die het programma wel, en die welke het programma niet gevolgd hadden. In vergelijking met de autochtone kinderen uit groep 5 in het normeringsonderzoek van de TAK-bovenbouw scoren de experimentescholen even hoog, en de leerlingen uit de controlegroepen laag tot zeer laag (Verhoeven & Vermeer, 1993, p. 53). Tussen de scores op de Tekstleestaak en de woordenschattoetsen in groep 1 en 2 bleken geen significante correlaties te zijn (Tekstleestaak $\times$ Passief $r=.16$, $\times$ Actief $r=.17$), maar wel in groep 4 (Tekstleestaak $\times$ Passief $r=.36$, $\times$ Actief $r=.31$, in beide gevallen $p<.01$). Op grond hiervan kan geconcludeerd worden dat ook de tweede vraag bevestigend beantwoord kan worden, namelijk dat het experimentele programma niet alleen op de woordenschat, maar ook op de begrip- en leesvaardigheid van de leerlingen een positief effect heeft gehad.

Zowel op de Passieve- als de Actieve-woordenschattoets scoren de controlescholen even hoog als Turken en Marokkanen uit het normeringsonderzoek van de TAK. De leerlingen van de experimentescholen daarentegen scoren op beide toetsen even hoog als Suri-naamse kinderen van die leeftijd. Het niveau van hun autochtone Nederlandse leeftijdsgenoten wordt echter niet bereikt. De receptieve woordenschat van Nederlandse kinderen van 8;3 jaar kan geschat worden op zo'n 5300 woorden, de produktieve woordenschat op

zo'n 3900 woorden (Verhoeven & Vermeer, 1992). Op basis van extrapolatie van de scores uit Tabel 1 neemt de omvang van de receptieve woordenschat van de experimentleerlingen toe van 1300 tot 4300 woorden, dat wil zeggen dat er een jaarlijkse groei is van ongeveer 800 à 900 woorden. Dat is minder dan de ruim 1000 die per jaar nodig was om hun achterstand in woordenschat in het Nederlands in te halen. Ook in produktieve woordenschat lopen de leerlingen van de experimentescholen, met een gemiddelde omvang van 3300 woorden, nog ongeveer een jaar achter bij hun Nederlandse leeftijdsgenoten. Anderzijds kan natuurlijk wel geconstateerd worden dat ze daarmee één à twee jaar voor lopen op hun leeftijdsgenoten van dezelfde etnische groep, zowel vergeleken met de landelijke normen, als de controlegroepen binnen en buiten de eigen school.

4 Conclusies en discussie

De resultaten van ons onderzoek laten zien dat het mogelijk is de woordenschat in het Nederlands van allochtone leerlingen een enorme impuls te geven. Terwijl de gehanteerde lesmaterialen niet speciaal voor dit doel gemaakt waren, en de tijdsinvestering van leerkracht en leerling niet buitengewoon groot was, nam de woordenschat van de leerlingen van de experimentescholen substantieel toe, meer dan die van de controlegroepen, en had die vergrote woordenschat ook een breder effect, namelijk op hun begrip- en leesvaardigheid in groep 5. Eind groep 4 hadden de leerlingen van de experimentescholen weliswaar niet het niveau bereikt van hun Nederlandse leeftijdsgenoten, maar liepen ze toch in woordenschat één à twee jaar voor op kinderen uit dezelfde etnische groep. Aangezien zij deze voorsprong ook hadden ten opzichte van de leerlingen van dezelfde experimenteschool, maar die niet het experimentele programma gevolgd hadden, zijn deze effecten toe te schrijven aan het experimentele programma.

Gegeven de resultaten van ons onderzoek denken we dat het mogelijk is onze derde onderzoeksvraag – namelijk om allochtone leerlingen de eerste vier schooljaren een zodanig onderwijsaanbod te geven dat ze ruim 1000 woorden per jaar leren en daarmee het niveau

van hun Nederlandse leeftijdsgenootjes halen – positief te beantwoorden, maar wel onder de volgende condities (zie ook Appel en Vermeer, 1996). Ten eerste moeten er in het aan de kinderen aangeboden programma elk jaar voldoende nieuwe woorden zitten, namelijk minstens 1200 woorden per jaar. Dat is ruim drie keer zoveel als in de huidige NT2-methoden aangeboden worden. Ten tweede moeten die woorden in zoveel mogelijk contexten (verhalen, prentenboeken, plaatjes, activiteiten) aangeboden en herhaald worden, om zo een goede verwerving te bevorderen. Ten derde moeten de leerlingen regelmatig, bijvoorbeeld vier dagen in de week elke dag een half uur, in een klein groepje van 4 à 6 leerlingen extra les krijgen, naast de klassikale aanbieding van de lesmaterialen en activiteiten. De verschillen in woordenschat tussen Nederlandse en allochtone leerlingen zijn namelijk zo groot, dat men moet differentiëren, ook al omdat tweetalige leerlingen vaak *andere* woorden moeten leren dan leerlingen die van huis uit Nederlands spreken (zie ook Appel & Vermeer, 1996). Alleen onder die voorwaarden bestaat de kans dat allochtone leerlingen hun achterstand in taalvaardigheid Nederlands inhalen, en daarmee hun positie op de onderwijsmarkt verbeteren.

Literatuur

- Aarnoutse, C.A.J., & Leeuwe, J.F.J. van (1988). Het belang van technisch lezen, woordenschat en ruimtelijke intelligentie voor begrip lezen. *Pedagogische Studiën*, 65, 49-59.
- Anderson, R.C., & Freebody, P. (1981). Vocabulary knowledge. In J.T. Guthrie (Ed.), *Comprehension and teaching: research reviews* (pp. 77-117). Newark, DE: International Reading Association.
- Appel, R., & Vermeer, A. (1990a). Uitbreiding woordenschat NT2 in de kleuterklassen. *Toegepaste Taalwetenschap in Artikelen*, 38 (3), 36-48.
- Appel, R., & Vermeer, A. (1990b). Woordkennis is sleutel voor schoolsucces. *Stimulans*, 8 (7), 6-8.
- Appel, R., & Vermeer, A. (1994). Met prentenboeken meer woordenschat. In P. Mooren & H. van Lierop (red.), *De kunst van het lezen* (pp. 157-172). Den Haag: NBLC.
- Appel, R., & Vermeer, A. (1996). Van oude methodes en verschillen die nog niet voorbij zijn: over NT1 en NT2 in het basisonderwijs. In S. Kroon & T. Valen (red.), *Het verschil voorbij. Onderwijs Nederlands als eerste en tweede taal in Nederland en Vlaanderen*. 's-Gravenhage: SDU.
- Boogaard, M., Damhuis, R., Gloppe, K. de, & Bergh, H. van den (1990). *De Nederlandse taalvaardigheid van allochtone en Nederlandse kleuters*. Amsterdam/Lisse: Swets & Zeitlinger.
- Karumma, P. (1992). Polysemie en tweede-taalverwerving. Amsterdam: intern rapport ATW, UvA.
- Kerkhoff, A. (1988). *Taalvaardigheid en schoolsucces*. Lisse: Swets & Zeitlinger.
- Mulder, L. (1994). *De derde fase van de OVB-cohort-onderzoeken in het basisonderwijs*. Nijmegen: ITS.
- Tesser, P.T.M. (1993). *Rapportage Minderheden 1993*. Rijswijk: Sociaal en Cultureel Planbureau.
- Tesser, P.T.M., Mulder, L., & Werf, G. van der (1991). *De eerste fase van de longitudinale OVB-onderzoeken. Het leerlingenonderzoek*. Nijmegen/Groningen: ITS/RION.
- Verhallen, M. (1994). *Lexicale vaardigheid van Turkse en Nederlandse kinderen*. Een vergelijkend onderzoek naar betekenis-toekenning. Amsterdam: Academisch proefschrift, Universiteit van Amsterdam.
- Verhoeven, L., & Vermeer, A. (1986) *Taaltoets Allochtone Kinderen*. Tilburg: Zwijssen.
- Verhoeven, L., & Vermeer, A. (1989). *Diagnose van kindertaal. Nederlandse taalvaardigheid van autochtone en allochtone kinderen*. Tilburg: Zwijssen.
- Verhoeven, L., & Vermeer, A. (1992). Woordenschat van leerlingen in het Basis- en MLK-onderwijs. *Pedagogische Studiën*, 69, 218-234.
- Verhoeven, L., & Vermeer, A. (1993). *Taaltoets Allochtone Kinderen Bovenbouw*. Tilburg: Zwijssen.
- Willet, J. (1987). Questions and answers in the measurement of change. *Review of research in education*, 15, 345-422.

Manuscript aanvaard 23-11-1995

Auteurs

R. Appel is bijzonder hoogleraar 'Verwerving en didactiek van het Nederlands als tweede taal' en als universitair hoofddocent verbonden aan de Vakgroep Algemene Taalwetenschap van de Universiteit van Amsterdam.

Adres: Spuistraat 210, 1012 VT Amsterdam.

A. Vermeer werkt als universitair docent bij het werkverband Taal & Minderheden van de Letterenfaculteit van de Katholieke Universiteit Brabant in Tilburg.

Adres: Postbus 90153, 5000 LE Tilburg.

Abstract

Speeding up the acquisition of Dutch vocabulary by migrant children

R. Appel & A. Vermeer. *Pedagogische Studiën*, 1996, 73, 82-92.

The importance of lexical skills in language development and school achievement is widely recognized. Migrant children in the Netherlands, however, lag far behind their Dutch classmates with respect to Dutch vocabulary. To speed up the acquisition of Dutch vocabulary by migrant children in the first four grades of primary school an experimental program was designed. In this article, the results of the study which evaluates this program are presented. In a longitudinal experiment, with an experimental and a control group ($n=57$, 53 resp.), both groups of children were pre-tested and post-tested each year with curriculum-dependent and -independent lexical tests. In addition, in grade 5 their text reading abilities were tested to examine whether the program had a broader effect. The outcomes of the experiment make clear that it is possible to substantially increase the Dutch vocabulary of migrant children. In each grade, the children of the experimental group performed the tasks much better than those of the control group. At the end of the experimental program, i.e. after four years, the children of the experimental group were one or two years ahead in their Dutch vocabulary compared to other second language learners of the same age group who did not participate in the experimental program.

Effecten van het verschil in bestuursvorm voor het openbaar en bijzonder basisonderwijs

R.H. Hofman

Samenvatting

In deze bijdrage wordt ingegaan op effecten van het verschil in bestuursvorm (publiekrechtelijk versus privaatrechtelijk) van besturen voor openbaar en voor bijzonder basisonderwijs. De contingentie-theoretische benadering van organisaties wordt daarbij als analysekader gebruikt. Er is nagegaan in hoeverre publiekrechtelijke en privaatrechtelijke schoolbesturen verschillen in hun bestuursorganisatie en in invloeden vanuit de lokale omgeving enerzijds, en in hoeverre dat samenhangt met verschillen in het bestuursbeleid dat zij ontplooiën anderzijds. De analyses betreffen een representatieve steekproef van 139 besturen – waarvan 44 met een publiekrechtelijke en 95 met een privaatrechtelijke bestuursvorm – voor openbaar en voor bijzonder basisonderwijs. De uitkomsten laten zien dat publiekrechtelijke en privaatrechtelijke schoolbesturen significant verschillen inzake de wijze waarop zij zijn georganiseerd en dat de invloed die de lokale omgeving op schoolbesturen uitoefent eveneens verschilt al naar gelang de bestuursvorm. Er wordt verder een verband tussen bestuursvorm en bestuursbeleid aangetoond: publiekrechtelijke schoolbesturen ontplooiën significant minder bestuursbeleid op twee van de vier maten van bestuursbeleid. Verschillende factoren treden op als predictoren van het bestuursbeleid van privaatrechtelijke en publiekrechtelijke schoolbesturen.

1 Inleiding en probleemstelling

Iedere onderwijsinstelling die uit de openbare kas wordt bekostigd, heeft een wettelijk erkend bevoegd gezag: het schoolbestuur. Dat bestuur is verantwoordelijk voor het functioneren van die school in al zijn facetten: bestuurlijk, financieel en onderwijskundig. Leune (1987) stelt

dan ook dat schoolbesturen in ons schoolwezen een bijzonder sterke machtspositie bekleden. Binnen de kaders die door het hoger gezag worden bepaald, zijn zij verantwoordelijk voor de totale gang van zaken op school. Het schoolbestuur is juridisch gezien het bevoegd gezag van een school en kan dan ook verantwoordelijk worden gesteld voor de kwaliteit van het onderwijs in die school. Deze verantwoordelijkheid verplicht de besturen activiteiten te ontplooiën ten aanzien van de onder het bestuur ressorterende scholen. Het bestuursbeleid in casu de activiteiten die een bestuur ontwikkelt, zijn ingekaderd door de regelgeving van de landelijke overheid. Binnen die bepalingen heeft een schoolbestuur een zekere mate van vrijheid in handelen: zijn beleidsvrijheid.

De andersoortige verantwoordelijkheid die besturen van openbaar en bijzonder onderwijs voor hun scholen dragen inzake de vrijheid van onderwijs zoals die formeel in de Grondwet is vastgelegd, kan het bestuursbeleid beïnvloeden. Bovendien kunnen uit het verschil in bestuursvorm, verschillen voortvloeien voor de organisatorische vormgeving van schoolbesturen. Het onderscheid tussen scholen voor openbaar en bijzonder onderwijs komt overeen met het onderscheid naar bestuursvorm. Openbare onderwijsinstellingen worden bestuurd door publiekrechtelijke organen, terwijl bijzonder scholen worden opgericht en bestuurd door particuliere rechtspersonen, zoals verenigingen en stichtingen.

In dit artikel worden de resultaten gepresenteerd van een onderzoek waarin besturen voor openbaar en voor bijzonder onderwijs beide zijn opgenomen. De probleemstelling richt zich op de bestuursvorm van die schoolbesturen en het effect daarvan op het bestuursbeleid. In het onderzoek wordt de hypothese getoetst dat de bestuursvorm van openbaar en bijzonder onderwijs als determinant optreedt van verschillen in bestuursbeleid. Deze vraag is van

belang in het lopende debat over de herkomst van effectiviteitsverschillen van scholen voor openbaar en voor bijzonder onderwijs. Als dergelijke effecten van bestuursvorm op bestuursbeleid kunnen worden vastgesteld kan nader worden onderzocht of de in Nederlandse studies regelmatig vastgestelde effectiviteitsverschillen tussen scholen voor openbaar en bijzonder onderwijs voor een deel daaraan kunnen worden toegeschreven.

De probleemstelling die in dit artikel wordt behandeld is als volgt geformuleerd:

Is er een verband aantoonbaar tussen de bestuursvorm (privaatrechtelijk versus publiekrechtelijk) van een schoolbestuur enerzijds en het bestuursbeleid van schoolbesturen anderzijds?

In de volgende paragraaf wordt ingegaan op de bestuursvorm in het Nederlandse onderwijsbestel. In paragraaf 3 worden de keuze van een geëigende organisatie-theoretische benadering voor analyse van het verschil in bestuursvorm en de invalshoeken die mogelijk zijn bij een analyse van de output van schoolbesturen, het bestuursbeleid, besproken. Opzet en methode van onderzoek worden besproken in paragraaf 4. De resultaten worden gepresenteerd in paragraaf 5. Daarbij wordt aandacht geschonken aan effecten van bestuursvorm op bestuursbeleid en predictoren van bestuursbeleid voor publiekrechtelijke en privaatrechtelijke besturen. Het artikel wordt afgesloten met conclusies en discussie van de resultaten in het licht van consequenties voor de bestuurlijke inrichting van ons onderwijsbestel.

2 De bestuursvorm in het Nederlandse onderwijsbestel

In het Nederlandse onderwijssysteem zijn, zoals eerder opgemerkt, belangrijke bevoegdheden toegekend aan de schoolbesturen; juridisch gezien zijn ze eindverantwoordelijke voor het onderwijs en de totale gang van zaken in de onder hun bestuur ressorterende scholen. Van Wieringen (1989) stelt dat de enorme verscheidenheid in vormgeving waarmee aan verantwoordelijkheden en bevoegdheden inhoud wordt gegeven, opvallend is.

In de bestuurlijke inrichting van ons onderwijsbestel staat centraal het verschil in verant-

woordelijkheid die besturen van openbaar en bijzonder onderwijs hebben inzake de vrijheid van onderwijs, zoals dat bij de pacificatie in 1917 formeel in de Grondwet is vastgelegd. Binnen het openbaar onderwijs is het verboden een specifieke godsdienst of levensbeschouwing tot uiting te brengen (artikel 23, lid 3 van de grondwet). Het openbaar onderwijs dient een pluriform karakter te hebben, terwijl het bijzonder onderwijs er juist voor dient om de eigen godsdienst of levensbeschouwing in het onderwijs tot uitdrukking te brengen.

Een eveneens belangrijk onderscheid tussen het openbaar en het bijzonder bestuur betreft de bestuursvorm (Akkermans & Leune, 1983). De formele verantwoording voor het openbare basisonderwijs ligt bij een publiekrechtelijk orgaan. De Wet op het basisonderwijs bepaalt dat het College van Burgemeester en Wethouders als bevoegd gezag van de lokale openbare school optreedt, voor zover de gemeenteraad niet anders bepaalt en, indien de gemeenteraad dit wenselijk oordeelt, met inachtneming van door deze raad te stellen regels (artikel 1). De formele verantwoordelijkheid ligt hier dus bij bezoldigde bestuurders. De meest voorkomende vorm van publiekrechtelijk beheer is in handen van de gemeente (i.c. het College van Burgemeester en Wethouders en de gemeenteraad). Andere publiekrechtelijke organen zoals functionele commissies of het bestuur krachtens de Wet Gemeenschappelijke Regelingen (WGR) komen veel minder voor. Recent is echter een ontwikkeling gaande waarin al meer voor dergelijke alternatieve publiekrechtelijke organen wordt gekozen. De formele verantwoording voor het bijzonder basisonderwijs ligt bij privaatrechtelijke organen, veelal stichtingen en verenigingen. Artikel 35 van de Wet op het basisonderwijs bepaalt dat een bijzondere school in stand wordt gehouden door een rechtspersoon met volledige bevoegdheid die zich blijkens de statuten of reglementen het geven van onderwijs ten doel stelt zonder daarbij het maken van winst te beogen. De leden van het schoolbestuur voor bijzonder onderwijs zijn in het basisonderwijs in het algemeen ouders die kinderen op de betreffende basisschool hebben zitten.

Schoolbesturen kunnen worden opgevat als vertegenwoordigers van groeperingen in de lokale gemeenschap. In het bijzonder confessio-

neel onderwijs is dit traditioneel (een deel van) de lokale geloofsgemeenschap. Voor het openbaar onderwijs ligt die vertegenwoordiging heel anders: de lokale verkiezingen bepalen daar wat verstaan moet worden onder de lokale gemeenschap. Van Wieringen (1989) stelt dat gemeenten in vergelijking met bijzondere schoolbesturen in een complexere positie zitten, omdat de cliëntencontroles of gemeenschapscontroles indirect verlopen waardoor aan het openbare schoolbestuur een politieke dimensie wordt gegeven. Anderzijds zal het bureaucratische karakter van het functioneren van het gemeentelijk apparaat de participatie van bij de school betrokken groeperingen beïnvloeden (vgl. Scott & Meyer, 1988; Van Wieringen, 1987; Van Wieringen, 1989).

3 Organisatie-theoretische benadering van schoolbesturen en het bestuursbeleid

Voor de analyse van de bestuurlijke inrichting van ons onderwijsbestel dient een daartoe geëigend organisatie-theoretisch kader te worden gehanteerd. De contingentie-theoretische benadering lijkt goed bruikbaar als algemeen conceptueel kader voor analyse van het functioneren van schoolbesturen (vgl. Lammers, 1987; De Leeuw, 1982). Alhoewel het contingentie-onderzoek behept is met speciale problemen zoals de (on)meetbaarheid van allerlei soorten van doelmatigheid en beperkte empirische precisie, is het theoretisch gezien relevant vanuit het gezichtspunt van bestuur- en beheersbaarheid van organisaties (Mintzberg, 1979; Lammers, 1987). Volgens Scheerens (1989) is de contingentie-benadering als paradigma voor organisatiekundige analyses dan ook niet meer weg te denken en ook zeer goed te rijmen met andere meta-theorieën over de effectiviteit van organisaties. Het gebruik van de contingentie-benadering voor organisatie-analyse van schoolbesturen in relatie tot de door hen bestuurde school biedt bovendien de mogelijkheid een brug te slaan tussen theorievorming omtrent organisatorische effectiviteit en schooleffectiviteit en de rol van schoolbesturen daarbij (Vgl. Kickert, 1979; De Leeuw, 1982).

In zijn contingentie-benadering van de school als systeem hanteert Scheerens (1990, p. 73) een systeem-analytisch model waarin vier clusters van relevante samenhangende factoren worden onderscheiden: context-, input-, proces- en outputvariabelen. Dit model houdt rekening met het feit dat het onderwijs bij uitstek een systeem is dat wordt gekenmerkt door een hiërarchische structuur: leerlingen behoren tot (sub)klassen, klassen vallen binnen scholen, scholen rekenen we tot schooltypen en tot hogere bestuurlijke lagen en zo verder (vgl. Barr & Dreeben, 1983; Hofman & Hofman, 1988). Binnen dat zogenaamde integrale model voor schooleffectiviteit wordt ervan uitgegaan dat niet sprake is van een directe relatie vanuit schoolbesturen tot de output van de school (leerlingprestaties) maar dat deze invloed via input en procesfactoren op schoolniveau en klasniveau tot het leerlingniveau loopt. De bijdrage van schoolbesturen valt binnen de context van scholen. Dit model houdt er rekening mee dat de schoolorganisatie een systeem van 'nested layers' is, (Barr & Dreeben, 1983) waarbij er rekening mee dient te worden gehouden dat elk organisatorisch niveau tegelijkertijd de context en de grenzen aangeeft voor het onderliggende niveau. Het hogere niveau zoals het schoolbestuur levert in deze zienswijze faciliterende condities voor de centrale processen op de lagere niveaus, het school- en klasniveau. Hofman, Hoebe en Guldmond (1995) analyseerden een onderzoeksmodel met als basis het contingentie-theoretische model van Scheerens en onderzochten het context-effect van schoolbesturen. De auteurs gingen er daarbij vanuit dat de invloed van de richting van een school (openbaar, protestants en katholiek) en de institutionele omgeving van de school, uitgedrukt in kenmerken van het schoolbestuur, samenhangt met de effectiviteit van scholen. De analyses laten zien dat de invloed van het schoolbestuur verloopt via procesvariabelen van de school.

Deze benadering van schoolbesturen als context-effect lijkt bovendien bij uitstek geschikt voor onderzoek waarbij schoolbesturen voor openbaar en voor bijzonder onderwijs worden onderzocht. Met name de invloed van omgevingskenmerken lijkt een belangrijke rol te spelen bij de analyse van functioneren en output van schoolbesturen voor bijzonder on-

derwijs versus schoolbesturen voor openbaar onderwijs. Scott en Meyer (1988) vertegenwoordigen de opvatting dat omgevingskenmerken verklarend zijn voor de interne situationele opbouw van het schoolbestuur. De auteurs passen deze opvatting toe op het verschil tussen openbaar en bijzonder onderwijs. Het openbaar onderwijs dient in beginsel de gehele gemeenschap, hetgeen betekent dat het in principe open staat voor verlangens van uiteenlopende groeperingen, met als mogelijk gevolg een zeker verlies aan interne samenhang. Bijzondere scholen richten zich op een deel van de lokale bevolking, waarbij de school zich rond die selectie organiseert en daarmee juist meer nadruk kan leggen op interne samenhang en samenbinding. Scott en Meyer (1988) constateren zowel op het beheersniveau als op het onderwijskundig niveau dat bepaalde omgevingskenmerken uiteenlopende gevolgen kunnen hebben voor het managementsgedrag van scholen en schoolbesturen. Deze analyse betreft de situatie in de Verenigde Staten; de hoofdprincipes van een inkadering van schoolbesturen in een publieke en private context zijn echter ook in Nederland te onderkennen (Van Wieringen, 1989, p. 150).

Er zijn verschillende invalshoeken en effectiviteitscriteria mogelijk om de output van een organisatie te beschrijven. In Figuur 1 zijn de voornaamste kenmerken van deze organisatie-theoretische opvattingen, het gehanteerde effectiviteitscriterium en het niveau waarop zich dat richt schématisch samengevat. Bovendien is aangegeven welke aandachtsgebieden voor analyse van het bestuursbeleid van schoolbesturen daaruit kunnen worden afgeleid. In on-

derhavig onderzoek ligt de nadruk op de vijf in Figuur 1 genoemde aandachtsgebieden van schoolbesturen. De frequentie en intensiteit waarmee schoolbesturen activiteiten ont-plooien op een of meerdere van die aandachtsgebieden en met name de mate waarin daarbij verschillen optreden tussen privaatrechtelijke en publiekrechtelijke besturen wordt onderzocht. In het onderzoek naar schoolbesturen is hiermee rekening gehouden bij de specificatie van het bestuursbeleid van schoolbesturen (vgl. Figuur 1: aandachtsgebieden). Een vijftal bestuursdimensies wordt in dit artikel onderscheiden, namelijk: onderwijsbeleid, personeelsbeleid, financieel-materieel beleid en de interne contacten en externe contacten van het schoolbestuur.

Het lijkt aannemelijk dat de eerder geschetste positie inzake de vrijheid van onderwijs van invloed is op de betrokkenheid van de schoolbesturen bij de inhoud van het onderwijs in de onder het bestuur ressorterende school. Uit onderzoek van Buurke (1988) kan worden opgemaakt dat de dubbelrol die de gemeente vervult ten aanzien van het onderwijs, namelijk die van lokale overheid voor het openbaar en voor het bijzonder onderwijs en die van schoolbestuur voor alleen het openbaar onderwijs in de gemeente, problemen oplevert. Buurke geeft dan ook als specifiek probleem voor het publiekrechtelijk bestuur aan: "Door de dubbele taakstelling betoont het gemeentebestuur zich in het algemeen voorzichtig in zijn schoolbestuurlijke attitude: het gemeentebestuur is weinig geneigd zich te identificeren met het openbaar onderwijs. In alle gevallen werden de

effectiviteitscriterium	niveau waarop gericht	aandachtsgebied schoolbesturen	theoretische achtergrond
produktiviteit	output en de determinanten	<i>onderwijskundig: primaire leerproces</i>	economische rationaliteit
aanpassingsvermogen	verwerven van inputs	<i>financieel/materieel</i>	organische systeemtheorie
betrokkenheid	motivatie	<i>personeelsbeleid</i>	'human relation' benadering
continuïteit	intern formele structuur	<i>relaties met interne geleidingen</i>	systeemleer, bureaucratie- en sociaal-psych. theoriën
responsiviteit	extern afhankelijkheid, macht	<i>relaties met externe groeperingen</i>	politieke theorie over organisaties

Bron: aangepast en uitgebreid naar Scheerens, 1989, p. 9

Figuur 1. Modellen van organisatie-theoretische effectiviteit en aandachtsgebieden voor schoolbesturen

nadelige gevolgen van de dubbele taakstelling onmiddellijk doorgespeeld aan de financiële gelijkstelling van openbaar en bijzonder onderwijs en de daarop gebaseerde doorbetalingsregeling. In het algemeen stelt het gemeentebestuur zich terughoudend op wat betreft het doen van (extra) uitgaven voor het openbaar onderwijs" (Buurke, 1988, p. 196). Onderzoek van Van der Venne en Van Wieringen (1995) toont eveneens belangrijke verschillen aan in het bestuurlijk functioneren van besturen voor openbaar en die voor bijzonder onderwijs. Bijzondere besturen zijn autonoom, vooral op financieel terrein, zij laten het beheren van budgetten vaker aan de school over. Het lijkt waarschijnlijk dat verschillen in bestuursbeleid tussen besturen voor openbaar en bijzonder onderwijs het meest pregnant tot uitdrukking zullen komen op bestuursterreinen die een sterk inhoudelijke component kennen waarbij die raakt aan de vrijheid van onderwijs. Dat lijkt met name het geval bij het door schoolbesturen gevoerde onderwijsbeleid en bestuursterreinen die direct financiële gevolgen hebben vanuit de doorberekeningssystematiek, zoals het financieel-materieel beleid en, zij het in mindere mate, het personeelsbeleid.

Overigens dient onderzoek naar het bestuursbeleid van schoolbesturen rekening te houden met een aantal kenmerken van de bestuursde school die het bestuursbeleid kunnen beïnvloeden. Hofman (1993) laat zien dat een tweetal kenmerken van de bestuursde school samenhangen met de activiteiten die schoolbesturen ontplooiën. Schoolbesturen die een grote school beheren ontplooiën meer onderwijsgericht bestuursbeleid, terwijl besturen die scholen beheren in geurbaniseerde gebieden meer personeelsgericht en minder financieel bestuursbeleid ontplooiën.

4 Methode van onderzoek: steekproef, kernbegrippen en toetsing

4.1 De steekproef

In het onderzoek zijn de schoolbesturen benaderd van een aselekt getrokken representatieve steekproef van Nederlandse basisscholen (Brandsma & Knuver, 1989). De respons van

de schoolbesturen bedroeg 65% voor het instrument waarmee het bestuursbeleid is gemeenten en 59% voor de organisatiekenmerken van besturen. De enquêtes zijn ingevuld door voorzitter en secretaris van het schoolbestuur van bijzondere scholen en door wethouder van onderwijs en het hoofd van de afdeling onderwijs in het geval van openbare scholen. In een vooronderzoek zijn de instrumenten beproefd en als betrouwbaar en valide beoordeeld. Non-responsonderzoek toonde aan dat de gegevens generaliseerbaar zijn naar de totale populatie van schoolbesturen van Nederlandse basisscholen (zie Hofman, 1993). In het onderzoek wordt gewerkt met een representatieve steekproef waarin 139 schoolbesturen van vier verschillende denominaties zijn opgenomen. In deze analyse wordt gebruik gemaakt van een tweedeling in besturen met een publiekrechtelijke bestuursvorm ($n=44$) en besturen met een privaatrechtelijke bestuursvorm ($n=95$). De publiekrechtelijke besturen kennen allen de integrale bestuursvorm, waarbij het College van B & W en de gemeenteraad gezamenlijk het bestuur over de openbare scholen uitoefenen.

4.2 De kernbegrippen

De kernbegrippen die centraal staan in deze studie zijn samengebracht in een ordening van vier groepen van factoren, namelijk (1) bestuursvorm, (2) contingentiefactoren ofwel factoren in de lokale omgeving van schoolbesturen, (3) structuurkenmerken van de bestuurlijke organisatie en (4) outputindicatoren, het bestuursbeleid van schoolbesturen. De gedachtengang hierbij is dat de bestuursvorm van schoolbesturen samenhangt met verschillen in de organisatorische structuur van schoolbesturen (verschillen in contingentie- of omgevingsfactoren) en via die factoren het bestuursbeleid beïnvloed.

Bij de *contingentiefactoren* gaat het om drie factoren uit de lokale omgeving van de onderzochte schoolbesturen. Allereerst de factor 'bestuurlijke schaal': hoeveel scholen ressorteren er onder het bestuur. Verder de factor 'invloed van schoolgeledingen en van externe instanties in de lokale omgeving van schoolbesturen op bestuursbeslissingen'. De besturen is gevraagd voor elke schoolgeleding (schoolleider, groepsleerkrachten, medezeggenschapsraad, ouderraad, niet-onderwijzend personeel, leer-

lingen, ouders in het algemeen e.d.) en ook voor organisatie en instanties die in de lokale omgeving van schoolbesturen opereren (onderwijsvakorganisaties, ondersteuningsinstellingen, inspectie, besturenorganisaties etc.) aan te geven hoe groot zij de invloed van die geleiding of instantie op bestuursbeslissingen acht. De derde factor betreft de mate waarin het schoolbestuur verschillende vormen van externe ondersteuning ontvangt (financiële, bouwkundige, onderwijskundige of betreffende personeelszaken).

Bij de *organisatiekenmerken* van schoolbesturen gaat het eveneens om drie factoren. De eerste factor betreft de interne bestuursopbouw: beschikt het bestuur over eigen interne ondersteuning via een bestuurs- of onderwijskundig bureau, en hoeveel interne organen kent het bestuur (werkgroepen, adviescommissies, inspraak-organen, etc.). De tweede factor betreft de deskundigheid en cohesie van het schoolbestuur. Vragen die daarbij zijn gesteld betreffen allereerst de mate waarin het bestuur beschikt over deskundigheid op een vijftal bestuursgebieden (personeels-beleid, financieel-economisch, materieel-bouwkundig, onderwijskundig en op levensbeschouwelijk terrein). In tweede instantie is gevraagd hoe groot het bestuur de cohesie tussen de bestuursleden inschat inzake het te volgen bestuursbeleid. De laatste factor betreft de mate waarin er binnen de besturen planmatig wordt gewerkt. Het gaat er daarbij om of er binnen het bestuur sprake is van: een document waarin het bestuursbeleid voor de langere termijn is vastgelegd; gespecificeerde functie-omschrijvingen voor de bestuursleden; duidelijke taakverdeling tussen de bestuursleden en een duidelijk taakafbakening van bevoegdheden van bestuur en school.

De psychometrische statistieken van de geconstrueerde schalen betreffende lokale omgevingsfactoren en organisatiekenmerken van schoolbesturen, zijn opgenomen in Bijlage 1 (Tabel 1a en 1b).

Het *bestuursbeleid* is gebaseerd op een (uitputtende) indeling van alle mogelijke bestuursactiviteiten. Op basis van beschikbare empirische gegevens is gekozen voor een indeling naar zes taakgebieden die de realiteit van het handelen van schoolbesturen het meest recht doet (vgl. Leune, 1981, 1987; CBKO, 1985; Drop, 1985; Detering & Kalkman, 1985; Van

Esch, 1988). Het gaat daarbij om de taakgebieden: onderwijsbeleid van schoolbesturen (16 items), organisatiebeleid (18 items), personeelsbeleid (11 items), financieel-materieelbeleid (13 items) en de interne en externe contacten van het bestuur. De betrouwbaarheid van de geconstrueerde schalen is voldoende tot hoog. Die zes taakgebieden representeren te zamen het bestuursbeleid. Om de vergelijkbaarheid voor openbaar en bijzonder onderwijs zo groot mogelijk te maken is de besturen gevraagd de totaal 78 deelactiviteiten (verdeeld over de zes taakgebieden) van bestuursbeleid te scoren voor één aselekt gekozen basisschool die zij besturen. Per taakgebied en deelelement dienden zij aan te geven hoe vaak per jaar zij een dergelijke activiteit ten uitvoer brachten (de frequentieschaal) en verder met welke soort van betrokkenheid dat het geval was, een meer flatterende of meer beleidsvormende (de taakopvattingsschaal). Een voorbeeld van de deelactiviteiten van het taakgebied onderwijsbeleid, zijn opgenomen in Bijlage 2, Tabel 2a.

De schoolbesturen blijken het meest actief te zijn op het terrein van de interne en externe contacten en op beheersmatig terreinen zoals het financieel-materieel beleid en het personeelsbeleid. De besturen zijn het minst actief op de onderwijsgerichte taakgebieden zoals het onderwijs- en organisatiebeleid. Dit komt overeen bevindingen uit ander besturen-onderzoek (Detering & Kalkman, 1985; Braster & Leune, 1986; Buurke, 1988; Van Esch, 1988).

Deze deeltaken zijn aan een factoranalyse onderworpen hetgeen resulteerde in vier bestuursdomeinen. In totaal wordt via deze factorstructuur van vier factoren 75% van de variantie in bestuursbeleid gebonden. Deze vier bestuursdomeinen representeren in het vervolg te zamen de criteriumvariabele bestuursbeleid namelijk (1) onderwijs-gericht bestuursbeleid, (2) communicatiegericht bestuursbeleid, (3) personeelsgericht bestuursbeleid en (4) financieel-materieel gericht bestuursbeleid (zie Bijlage 2, Tabel 2b). Voor meer informatie omtrent specifieke items wordt verwezen naar Hofman (1993).

4.3 De toetsing

In de volgende paragraaf worden de uitkomsten besproken van analyses van de samenhang tussen bestuursvorm enerzijds en het bestuurs-

beleid van schoolbesturen anderzijds. Vervolgens is via regressie-analyses nagegaan welke van de contingentie- en structuurfactoren als determinanten van bestuursbeleid optreden en in hoeverre dat verschilt naar bestuursvorm. Omdat uit ander besturen-onderzoek is gebleken dat een aantal variabelen op schoolniveau, zoals schoolgrootte, urbanisatiegraad en in hoeverre sprake is van een stimuleringschool, van invloed zijn op het ontplooiende bestuursbeleid is daarvoor in de analyses gecontroleerd.

5 Resultaten

5.1 Effecten van bestuursvorm op bestuursbeleid

Onderzoek laat zien dat publiekrechtelijke en privaatrechtelijke besturen onderling nogal verschillen in de invloed die de lokale omgeving uitoefent op hun functioneren en evenzeer wat betreft hun bestuursorganisatie (Van Esch, 1988; Van Wieringen, 1989; Hofman, 1993). In dit artikel ligt het accent op de vraag in hoeverre verschillen in bestuursvorm samenhangen met verschillen in bestuursbeleid dat die besturen ontplooiën. Per bestuursvorm worden in Tabel 1 de gemiddelde scores (factorscores) van bestuursbeleid (uiteengelegd in een viertal bestuursdomeinen) gepresenteerd. Bij inspectie van Tabel 1 is een lichte trend waarneembaar, de publiekrechtelijke schoolbesturen zijn op alle vier de beleidsdomeinen minder actief dan de privaatrechtelijke besturen. Van significante verschillen is sprake op twee van de vier bestuursdimensies. Publiekrechtelijke besturen voor openbaar onderwijs onderhouden minder frequent en intensief contacten met interne geledingen in de school en met externe instanties ten behoeve van de bestuursde school, dan de privaatrechtelijke schoolbesturen. Bij het financieel/materieel bestuursdomein is eveneens sprake van een significant verschil

tussen publiekrechtelijke schoolbesturen enerzijds en privaatrechtelijke schoolbesturen anderzijds. Deze uitkomst bevestigt de eerder uitgesproken verwachting dat het verschil in opdracht van publiekrechtelijke en privaatrechtelijke schoolbesturen enerzijds en de financiële gelijkstelling van het openbaar en bijzonder onderwijs die gestalte krijgt in de zogenaamde overschrijdingsregeling anderzijds, van invloed is op het ten uitvoer gebrachte bestuursbeleid op het terrein van onderwijsbeleid en op financieel-materieel terrein.

5.2 Predictoren van bestuursbeleid

Vervolgens zijn afzonderlijke regressie-analyses uitgevoerd voor de publiekrechtelijke en privaatrechtelijke schoolbesturen voor openbaar en bijzonder onderwijs. Met deze separate analyses kan worden nagegaan of het verschil in contingentie - in casu omgevingsfactoren en organisatiekenmerken van publiekrechtelijke en privaatrechtelijke besturen - wellicht eveneens een verschillende betekenis hebben in het licht van komen tot uitvoering van de vier onderscheiden soorten van bestuursbeleid. Bij de analyses worden telkens eerst de drie controlevariabelen (schoolgrootte, urbanisatiegraad en stimuleringschool) in de analyse opgenomen. Om te voorkomen dat verschillen in bestuursbeleid aan verschillen op die controlevariabelen worden toegeschreven, worden zij als eerste stap in de regressie-analyses opgenomen.

Wanneer we Tabel 2 inspecteren met het oog op overeenkomsten en verschillen tussen beide subsets van schoolbesturen dan is allereerst opmerkelijk dat de variantie die wordt gebonden door de opgenomen contingentie- of omgevingsfactoren en organisatiekenmerken per bestuursvorm nogal verschilt. De verklaarde variantie voor de publiekrechtelijke besturen

Tabel 1
Bestuursbeleidsscores naar bestuursvorm

Bestuursvorm	onderwijs-gericht bestuursbeleid	communicatie-gericht bestuursbeleid	personeel-gericht bestuursbeleid	financieel-materieel beleid
Publiekrechtelijk	-.06	-.24*	-.02	-.68**
Privaatrechtelijk	.03	.11	.06	.30

(** = $p < .05$; * = $p < .01$)

Tabel 2

Predictoren van bestuursbeleid van publiekrechtelijke en privaatrechtelijke besturen op de vier bestuursdomeinen (verklaarde variantie en bèta)

A. Onderwijsgericht bestuursdomein	Publiekrechtelijk		Privaatrechtelijk	
	r^2	bèta	r^2	bèta
1. Covariaten:	-	-	-	-
- urbanisatiegraad	.45	.67		
2. Contingentiefactoren:				
- invloed interne geledingen	.54	.33	.16	.40
- breedte externe ondersteuning			.20	.22
3. Structuurkenmerken:				
- planmatigheid bestuur	.61	.33	-	-
B. Communicatief domein	Publiekrechtelijk		Privaatrechtelijk	
	r^2	bèta	r^2	bèta
1. Covariaten:	-	-	-	-
2. Contingentiefactoren:	-	-		
- invloed externe groeperingen			.09	.30
3. Structuurkenmerken:	-	-		
- interne bestuursdifferentiatie			.15	.24
- vergaderstructuur			.20	.23
C. Personeelsgericht domein	Publiekrechtelijk		Privaatrechtelijk	
	r^2	bèta	r^2	bèta
1. Covariaten:	-	-		
- urbanisatiegraad			.13	.36
2. Contingentiefactoren:			-	-
- invloed externe groeperingen	.24	.49		
3. Structuurkenmerken:				
- breedte bestuursdeskundigheid	.33	.31		
- planmatigheid bestuur			.19	.25
D. Financieel-materieelgericht domein	Publiekrechtelijk		Privaatrechtelijk	
	r^2	bèta	r^2	bèta
1. Covariaten:	-	-		
- urbanisatiegraad			.12	-.34
2. Contingentiefactoren:				
- aantal externe ondersteuningsinstanties	.20	.44		
- vergaderfrequentie met reguliere geledingen			.17	-.22
3. Structuurkenmerken:				
- interne bestuursdifferentiatie	.31	.34		
- aantal interne organen			.24	-.27
- planmatigheid bestuur			.28	.22

varieert van 0% (communicatief domein) tot maximaal 61% (onderwijsgericht domein). Voor de privaatrechtelijke besturen lopen de percentages van 19% (personeel) tot maximaal 28% (financieel-materieel). Bovendien valt op dat relevante predictoren van bestuursbeleid nogal verschillen per bestuursvorm. Hieronder wordt in het kort gerecapituleerd welke factoren uit de verschillende clusters van publiekrechtelijke en privaatrechtelijke schoolbesturen optreden als determinanten van bestuursactiviteit.

Het publiekrechtelijk bestuur van een school

voor openbaar onderwijs ontplooit relatief meer onderwijsgerichte bestuursactiviteit indien sprake is van:

- een onder het bestuur ressorterende school in een (grote) stad; waarbij
- de interne geledingen relatief veel invloed uitoefenen op de bestuursbeslissingen; en bovendien
- het handelen van het bestuur als planmatig is te kenschetsen.

Dit bestuur ontplooit relatief meer personeelsgerichte bestuursactiviteit indien sprake is van:

- een relatief sterke invloed van externe groeperingen op bestuursbeslissingen; en

b het bestuur zichzelf inschat als een bestuur met deskundigheid op een breed terrein. Dit bestuur ontplooit relatief meer *financieel-materiële* bestuursactiviteit indien sprake is van:

- a ondersteuning van relatief veel verschillende ondersteuningsinstanties; en
- b een hoge mate van interne bestuursdifferentiatie.

Het *privaatrechtelijk bestuur van een school voor bijzonder onderwijs* ontplooit relatief meer onderwijsgerichte bestuursactiviteit indien sprake is van:

- a een grote invloed van interne geledingen op de bestuursbeslissingen; en
- b een brede externe ondersteuning op meerdere voor schoolbesturen relevante deskundigheidssterreinen.

Dit bestuur ontplooit relatief meer *communicatieve* bestuursactiviteit indien sprake is van:

- a een relatief sterke invloed van externe instanties op bestuursbeslissingen;
- b een relatief hoge mate van interne bestuursdifferentiatie; en
- c bovendien een vergaderstructuur met planmatige besluitvorming.

Dit bestuur ontplooit relatief meer *personeelsgerichte* bestuursactiviteit indien sprake is van:

- a een school in een (grote) stad; alsmede van
- b een schoolbestuur waarvan het bestuursmatig handelen als planmatig kan worden gekenschetst.

Dit bestuur ontplooit relatief meer *financieel-materieelgerichte* bestuursactiviteit indien sprake is van:

- a een school op het platteland; waarbij
- b weinig overleg tussen schoolbestuur en schoolgeledingen; en
- c binnen het schoolbestuur geen interne organen zoals werkgroepen, adviesorganen en werkgroepen aanwezig zijn; terwijl bovendien
- d het handelen van het schoolbestuur als planmatig kan worden beoordeeld.

Voorgaande maakt duidelijk dat er wel sprake is van overlap tussen de predictoren van bestuursbeleid voor publiekrechtelijke en privaatrechtelijke schoolbesturen maar dat zij niet per bestuursdomein dezelfde rol vervullen. De bestuursactiviteit van publiekrechtelijke en

privaatrechtelijke besturen wordt alleen in het geval van het onderwijsgerichte bestuursbeleid door eenzelfde predictor verklaard. De factor 'invloed van interne geledingen in en rondom de school op bestuursbeslissingen' hangt positief samen met de mate waarin (publiekrechtelijke en privaatrechtelijke) besturen onderwijsgericht bestuursbeleid ontplooiën.

6 Discussie: de bestuurlijke inrichting van ons onderwijsbestel

In de probleemstelling stond de vraag centraal of er een verband aantoonbaar is tussen de bestuursvorm (privaatrechtelijk versus publiekrechtelijk) en het bestuursbeleid van schoolbesturen anderzijds. De bevindingen van de analyses tonen aan dat een dergelijke relatie bestaat. Wanneer we het bestuursbeleid beschouwen, blijkt sprake van een licht contrast tussen schoolbesturen voor openbaar en voor bijzonder onderwijs. Publiekrechtelijke besturen voor openbaar onderwijs ontplooiën op alle vier de onderscheiden bestuursdomeinen beneden het gemiddelde, terwijl de privaatrechtelijke op alle vier boven het gemiddelde bestuursbeleid ontplooiën. Een significant verschil is echter alleen vastgesteld voor twee van de vier bestuursdomeinen namelijk het communicatiegerichte bestuursbeleid en het financieel-materieel bestuursbeleid. Het feit dat minder overleg wordt gevoerd met de direct betrokkenen van de openbare scholen kan worden verklaard door het verschil in bestuurlijke schaal. Een grootschalig schoolbestuur heeft minder mogelijkheden om zich intensief bezig te houden met het reilen en zeilen van die ene specifieke school. Het bestuursbeleid zal dan ook weinig schoolspecifiek zijn.

Voor het geringere bestuursbeleid van de publiekrechtelijke besturen voor openbaar onderwijs op financieel-materieel terrein kan een verklaring worden gezocht in de zogenaamde 'overschrijdingsregeling'. Dit houdt in dat, wanneer de gemeente als schoolbestuur besluit bepaalde fondsen in te zetten voor het openbaar onderwijs, men diezelfde fondsen ook aan het bijzonder onderwijs dient uit te betalen. In de praktijk zal een gemeente die zijn begroting

rond wil houden dan ook terughoudend opereren en niet snel besluiten om extra subsidies, leermaterialen of anderszins extra zaken aan de openbare scholen te geven.

Voorspellers van bestuursbeleid

Het bestuursbeleid van publiekrechtelijke en privaatrechtelijke besturen wordt alleen in het geval van het onderwijsgerichte bestuursdomein door eenzelfde predictor verklaard. Het gaat daarbij om de specifieke gerichtheid van het bestuur op de unieke kenmerken van elke onder het bestuur ressorterende school. Dat wil zeggen dat bij schoolbesturen ongeacht hun bestuursvorm de invloed die de verschillende interne geledingen in de schoolorganisatie zoals de schoolleider, groepsleerkrachten en ouders uitoefenen op beslissingen van het schoolbestuur samenhangt met de mate waarin men onderwijskundige beslissingen neemt voor de school. Alhoewel in deze studie empirische regelmatigheid zijn vastgesteld blijft nog onduidelijk welke processen ten grondslag liggen aan de geconstateerde verbanden. Een argumentatie dat schoolbesturen die de interne geledingen van de bestuurde school veel invloed toekennen, daarmee een bestuursbeleid ontwikkelen dat gedragen wordt door de hele school staat tegenover de mogelijkheid dat de bestuursinvloed van de interne geledingen vooral betekent dat niet het bestuur bestuurt, maar de school(leider).

Relevante voorspellers van bestuursbeleid zijn verder de mate waarin men externe ondersteuning ontvangt en de mate waarin sprake is van een sterk geïnstrumenteerd en gestructureerd bestuur. De instrumentele kant van het bestuur verwijst naar een effectief schoolbestuur dat breed deskundig is. Een dergelijk bestuur zou moeten bestaan uit leden met verschillende deskundigheden zoals kennis van personeelszaken, van financieel- materiële en bouwkundige zaken, van de onderwijswetgeving en zeker ook onderwijskundige deskundigheid. Bij een gestructureerd bestuur is sprake van een bestuur dat werkt met duidelijke naar inhoud en taak gespecificeerde functieomschrijvingen, een heldere taakverdeling tussen bestuursleden en een eenduidige afbakening van bevoegdheden van schoolbestuur en schoolleiding. Een gestructureerd bestuur maakt tijdens haar vergaderingen gebruik van

een agenda, van verslaglegging, van een besluitenlijst en controle op de uitvoering ervan. Bij zo'n bestuur is bovendien sprake van een document waarin het bestuursbeleid voor de langere termijn is vastgelegd. Voornoemde factoren hangen bij publiekrechtelijke en privaatrechtelijke besturen samen met de mate waarin zij verschillende vormen van bestuursbeleid ontplooiën.

Wenselijke bestuursvorm

Voorgaande betekent dat de contingentie-theoretische benadering van het schoolbestuur als organisatie de empirische werkelijkheid recht doet. De andersoortige verantwoordelijkheid die besturen voor openbaar en bijzonder onderwijs hebben inzake de vrijheid van onderwijs, zoals die bij de pacificatie in 1917 formeel in de Grondwet is vastgelegd, komt tot uitdrukking in het verschil in bestuursvorm: de privaatrechtelijke voor het bijzonder onderwijs versus de publiekrechtelijke voor het openbaar onderwijs. Alhoewel de discussie over de wenselijke bestuursvorm voor het openbaar onderwijs in laatste instantie een politieke beslissing is en niet langs empirische weg tot een conclusie is te voeren, geven de resultaten van dit onderzoek aanleiding tot een keuze van bestuursvorm waarbij (a) de 'overschrijdingsregeling' wordt opgeheven en waarbij (b) de geledingen in en rond de school de beslissingen van het bestuur kunnen beïnvloeden. Dit betekent in feite dat er een bestuur noodzakelijk is dat dicht bij die ene specifieke basisschool staat. Ook in de Verenigde Staten wordt veel aandacht besteed aan de wenselijke bestuursvorm voor het openbaar 'public' onderwijs (Coleman, Hoffer & Kilgoré, 1985; Hofman, in press). Chubb en Moe (1990) stellen bijvoorbeeld vast dat meer dan 60% van de effectieve scholen beschikken over meer dan gemiddelde bestuurlijke autonomie, tegen slechts een derde deel van de niet-effectieve scholen. Hannaway (1991) stelt vast dat de institutionele omgeving van openbare ('public') en katholieke scholen samenhangt met de professionele attitude en het werkgedrag van leerkrachten en zo de educatieve leeromgeving van leerlingen beïnvloedt. Deze bevindingen sluiten aan bij het benadrukken van het belang van een bestuursorganisatie en bestuursbeleid waarin het reilen en zeilen en de wensen van de geledingen van die ene unieke

school, centraal staan. De bestuursvorm die daarbij wordt gewenst betreft een loskoppeling van het bestuur van de openbare school van de publiekrechtelijke overheidssector.

Ook in Nederland komen wijzigingsvoorstellen omtrent de bestuursvorm voor het openbaar onderwijs meestal neer op een ont koppeling van de rol van schoolbestuur en van de algemene verantwoordelijkheid van de gemeentelijke overheid voor het gehele (lokale) onderwijs. Vele mogelijkheden zijn naar voren gebracht dan wel in de praktijk gerealiseerd zoals de functionele commissie, de samenwerkingsschool, de privaatrechtelijke stichting of vereniging met het statutair vastgelegd openbaar karakter, de nationale school, het onderwijsschap en dergelijke. De laatste tijd komt het al vaker voor dat de gemeenteraad bij verordening het bevoegd gezag over een openbare school delegeert aan een commissie krachtens artikel 61 van de Gemeentewet (vgl. Leune, 1987, p. 253; Braster, Cohen, Dodde & Verkruijsen, 1991). Uit cijfers van het ITS blijkt dat in 1992 zo'n 12% van alle gemeenten zo'n bestuurscommissie voor haar openbaar basis- of voortgezet onderwijs heeft ingesteld. De bestuurscommissie wordt wel gezien als een semi-publiekrechtelijke beheersvorm. Uit onderzoek van Van Kessel (1993) komt naar voren dat de door commissies bestuurde scholen positiever zijn over hun bestuur dan scholen die rechtstreeks door B & W worden bestuurd. Zij achten de kwaliteit van het bestuur beter, de afstand tussen bevoegd gezag en ouders kleiner en de belangenbehartiging van het openbaar onderwijs beter. Ook geven zij aan dat de aanwezigheid van (onderwijs-) deskundige leden belangrijk is. Uit onderzoek van Hofman en Vonkeman (1993) onder besturen en directies voor openbaar voortgezet onderwijs blijkt bij de vraag welke bestuursvorm in principe de belangen van de openbare school het beste kan dienen, dat de integrale publiekrechtelijke bestuursvorm minder vaak wordt genoemd dan de bestuurscommissie ex art. 61.

De bevindingen beschreven in dit artikel bevestigen dat de institutionele omgeving en de bestuurlijke inrichting van publiekrechtelijke en privaatrechtelijke schoolbesturen nogal verschilt. Overigens maakt de invloed van de bestuurlijke schaal van besturen het lastig uitspreken te doen over het schoolbestuur voor

openbaar of bijzonder onderwijs (Hofman, 1994; Van Wieringen & Biemond, 1991). In slechts 21% van de Nederlandse gemeenten met een openbare school is bijvoorbeeld sprake van een zelfstandige afdeling Onderwijs. In kleine gemeenten wordt een deel van het bestuurlijk werk dan ook veelal door de schoolleiding uitgevoerd. In de praktijk blijken alleen grote gemeentebesturen te beschikken over een eigen bestuursondersteunend apparaat dat op het openbaar en het bijzonder onderwijs is georiënteerd. De privaatrechtelijke besturen voor het bijzonder onderwijs vervullen echter daarentegen alleen een functie voor de scholen onder beheer.

In het lopende debat over de herkomst van effectiviteitsverschillen van scholen voor openbaar en voor bijzonder onderwijs (vgl. Hofman, Hoebe & Guldmond, 1995; Hofman, Hofman, Guldmond & Dijkstra, 1995), wordt het interessant na te gaan in hoeverre dergelijke verschillen in institutionele omgeving, als uitvloeisel van de bestuurlijke inrichting c.q. de bestuursvorm, een gedeeltelijke verklaring bieden voor de in Nederlandse studies regelmatig vastgestelde effectiviteitsverschillen tussen scholen voor openbaar en bijzonder onderwijs.

Literatuur

- Akkermans, P.W.C., & Leune, J.M.G. (red.). (1983). *Het bestuur van het openbaar onderwijs*. Den Bosch: Malmberg.
- Barr, R., & Dreeben, R. (1983). *How schools work*. Chicago: University of Chicago Press.
- Brandsma, H.P., & Knuver, J.W.M. (1989). *Basis-schoolkenmerken als determinanten van het functioneren van leerlingen*. Deel I: instrumentconstructie. Deel II: Resultaten. Groningen: RION.
- Braster, J.F.A., Cohen, M.J., Dodde, N.L., & Verkruijsen, W.G. (1991). *De bestuursvorm van het openbaar onderwijs: knelpunten en alternatieven*. Almere: Vereniging van Openbaar Onderwijs.
- Braster, J.F.A., & Leune, J.M.G. (1986). *Samenstelling en functioneren van schoolbesturen*. Rotterdam: Erasmus Universiteit, Vakgroep Onderwijssociologie en Onderwijsbeleid.
- Buurke, W.J. (1988). *Het gemeentebestuur als lokale overheid en als schoolbestuur*. Nijmegen: ITS.

- CBKO (1985). *Het besturen van scholen*. Discussie-nota. 's-Gravenhage.
- Chubb, J.E., & Moe, T.M. (1990). *Politics, markets and America's schools*. Washington, DC: The Brookings Institutions.
- Coleman, J.S., Hoffer, T., & Kilgore, S. (1985). Achievement and segregation in secondary schools: a further look at public and private school differences. *Sociology of Education*, 55, 162-182.
- Detering, P., & Kalkman, W. (1985). *Besturen met beleid*. Amsterdam: Universiteit van Amsterdam, Vakgroep Onderwijskunde.
- Drop, H. (1985). *Algemene inleiding onderwijsrecht*. Zwolle: Tjeenk Willink.
- Esch, W. van (1988). *Schoolbesturen en de vrijheid van onderwijs*. Academisch proefschrift. Nijmegen: ITS.
- Hannaway, J. (1991). The organization and management of public and catholic schools: looking inside the 'black box'. In E. Jemenez & M.E. Lockheed (Eds.), *Private versus public education an international perspective*. *International Journal of Educational Research*, 15, 463-483.
- Hofman, R.H. (1993). *Effectief schoolbestuur. De bijdrage van schoolbesturen aan de effectiviteit van basisscholen*. Academisch proefschrift. Monografieën onderwijsonderzoek 13. Groningen: RION/RUG.
- Hofman, R.H. (1994). Schoolbesturen en effectiviteit van basisscholen. In B.P.M. Creemers (Red.), *Dereguleren en de kwaliteit van het onderwijs*. Monografieën onderwijsonderzoek 17 (pp. 89-101). Groningen: RION/RUG.
- Hofman, R.H. (in press). Contextual influences on school effectiveness: the role of school boards. *School effectiveness and school improvement*.
- Hofman, R., Hoeben, W., & Guldmond, H. (1995) Denominatie en effectiviteit van schoolbesturen. *Tijdschrift voor Onderwijsresearch*, 20, 63-78.
- Hofman, R.H., & Hofman, W.H.A. (1988). Definiering van onderwijseffectiviteit: kanttekeningen vanuit het onderwijsonderzoek. In B.P.M. Creemers (Red.), *Handboek voor schoolorganisatie en onderwijsmanagement* (6310-1 t/m 6310-17). Alphen aan den Rijn: Samsom.
- Hofman, R.H., Hofman, W.H.A., Guldmond, H., & Dijkstra, A.B. (1995). *Variation in effectiveness in private and public schools: the impact of functional communities*. Artikel voor de AERA te San Francisco, USA.
- Hofman, R.H., & Vonkeman, E.B. (1993). *Het openbaar onderwijs in fusie-operaties*. Groningen: RION/RUG.
- Kessel, N. van (1993). *Integraal of alternatief. Commissies ex artikel 61 en gemeenschappelijke regelingen als bestuursvorm van het openbaar onderwijs*. Nijmegen: ITS.
- Kickert, W.H.M. (1979). *Organization of decision making, a systems-theoretical approach*. Amsterdam: North Holland Publishing Company.
- Lammers, C.J. (1987). *Organisaties vergelijkend onderwijs*. Utrecht/Antwerpen: Het Spectrum.
- Leeuw, A.C.J. de (1982). *Organisaties: management, analyse ontwerp en verandering*. Assen: Van Gorcum.
- Leune, J.M.G. (1981). De administratieve structuur van het Nederlandse schoolwezen. In J.A. van Kemenade (Red.), *Onderwijs: bestel en beleid* (pp. 333-374). Groningen: Wolters-Noordhoff.
- Leune, J.M.G. (1987). Besluitvorming en machtsverhoudingen in het Nederlandse onderwijsbestel. In J.A. van Kemenade e.a. (Red.), *Onderwijs: bestel en beleid 2. Onderwijs en samenleving*. B, (pp. 238-436). Groningen: Wolters-Noordhoff.
- Mintzberg, H. (1979). *The structuring of organizations: a synthesis of the research*. Englewood Cliffs, NJ: Prentice Hall.
- Scheerens, J. (1989). *Wat maakt scholen effectief: samenvatting en analyse van onderzoeksresultaten*. Den Haag: SVO. (Balans van onderzoek).
- Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School effectiveness and school improvement*, 1(1), 61-80.
- Scott, W.R., & Meyer, J.W. (1988). Environmental linkages and organizational complexity: public and private schools. In Th. James & H.M. Levin, *Comparing public and private schools* (pp. 128-160). New York: Falmer Press.
- Venne, L.H.J. van der, & Wieringen, A.M.L. van (1995). *Schoolbesturen in het bijzonder onderwijs. Veranderingen in de periode 1991-1994*. Amsterdam: SCO-Kohnstamm Instituut.
- Wieringen, A.M.L. van (1987). Schoolbestuurlijke toerusting en oriëntatie. *Handboek Schoolorganisatie en Onderwijsmanagement*.
- Wieringen, A.M.L. van. (1989). *Bestuur en management van onderwijsinstellingen*. Groningen: Wolters-Noordhoff.
- Wieringen, A.M.L. van, & Biemond, R. (1991). *Gemeentelijk onderwijsbeleid en schoolbestuur in hoofdlijnen*. Amsterdam: Universiteit van Amsterdam, Vakgroep Onderwijskunde.

Auteur

R.H. Hofman is als senior-onderzoeker verbonden aan het Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling (GION) van de Rijksuniversiteit van Groningen.

Adres: Westerhaven 15, 9718 AW Groningen.

Abstract

Effects of differences in form of governance in public and private primary education

R.H. Hofman. Pedagogische Studiën, 1996, 73, 93-107.

Based on data of a group of 139 school boards of primary schools effects of differences in the form of governance of private and public schools were examined. Differences were found for the organizational features of these school boards and for the influence of the local community on the policy of the school boards. The results show further a significant difference in administrative control for school boards of public and private schools on two dimensions: finance and communication. The administrative control on finance is not as strong in public as it is in private schools. Furthermore, school boards of private schools communicate more often with the members of their school team than school boards of public schools. Predictors of administrative control vary between school boards of public and private education.

Bijlage 1

Contingentie-factoren en organisatie-kenmerken van schoolbesturen

Tabel 1a

Psychometrische karakteristieken van drie contingentiefactoren (gemiddelde, standaard deviatie, aantal items, bereik en Cronbach's alpha)

	X	(SD)	items	range	α
(1) Bestuurlijke schaal:					
– aantal basisscholen	6.7	(8.7)	1	1-44	nvt
(2) Invloed lokale omgeving:					
– invloed van reguliere geledingen op bestuursbeslissingen	2.0	(0.5)	8	1- 4	.73
– invloed van externe instanties (zoals ondersteuningsinstellingen, koepels, inspectie, lokale overheid, etc.) op bestuursbeslissingen	1.9	(0.5)	13	1- 4	.77
(3) Externe ondersteuning:					
– breedte van externe ondersteuning	2.0	(0.5)	5	0- 5	.73

Tabel 1b

Structuurkenmerken van schoolbesturen

Structuurkenmerken	X	(SD)	items	range	α
(1) Bestuursdifferentiatie					
– interne bestuursopbouw	3.7	(1.1)	3	2- 6	.80
(2) Deskundigheid en cohesie					
– deskundigheid	3.2	(0.5)	4	1- 4	.68
– cohesie	3.4	(0.6)	1	1- 4	nvt
(3) Planmatigheid					
– planmatigheid bestuur	11.6	(1.0)	3	4-12	.72

Bijlage 2

Voorbeeld-items taakgebied onderwijsbeleid en factorstructuur taakgebieden

Tabel 2a

Onderwijsbeleid: betreft de grondslag, doelen, middelen en evaluatie van het onderwijs op school:

- levensbeschouwelijke en/of
- godsdienstige identiteit;
- pedagogisch-didactische grondslag;
- concrete leerdoelen;
- onderwijskundige vernieuwingen;
- onderwijsinhoudelijke gerichtheid van de schoolleiding;
- verdeling leertijd (o.a. lesrooster);
- leerdoelgerichtheid teamleden;
- schoolwerkplan;
- activiteitenplan;
- keuze leermiddelen;
- didactische werk/groeperingsvormen;
- leer-/en prestatieklimaat;
- evaluatie onderwijsdoelen;
- leerlingvoortgangsregistratie;
- aansluiting basis- voortgezet onderwijs;
- gedragsregels voor leerlingen.

Tabel 2b

Factorstructuur taakgebieden

Factoren	1	2	3	4
(1) Onderwijsbeleid	.86			
OBFREQ				
OBTAAK	.77			
(2) Organisatiebeleid	.80			
ORFREQ				
ORTAAK	.71			
(3) Interne contacten		.75		
ICFREQ				
ICTAAK		.75	.52	
(4) Externe contacten		.67		
ECFREQ				
ECTAAK		.68		
(5) Financieel/materieel beleid			.70	
FMFREQ				
FMTAAK			.86	
(6) Personeelsbeleid				.81
PRFREQ				
PRTAAK				.83
Eigenwaarde	5.41	1.41	1.14	1.00
Verklaarde variantie	45.1%	11.8%	9.5%	8.4%

(Toelichting: alleen ladingen van >.40)

Een bijdrage tot de validatie van het (meta)cognitieve deel van de Inventaris Leerstijlen

F. J. Prins, V. V. Busato, C. Hamaker en K. H. Visser

Samenvatting

Met de Inventaris Leerstijlen (ILS) van Vermunt (1992), een zelfrapportage vragenlijst, is het mogelijk van studenten een leerstijlprofiel samen te stellen. Men kan zich echter afvragen of studenten studeren zoals ze hebben aangegeven met de ILS. Het doel van dit onderzoek is een bijdrage te leveren tot de validatie van een deel van de ILS.

Zevenentwintig eerstejaars psychologie-studenten, vijf met een reproductiegericht, acht met een betekenisgericht en dertien met een ongericht leerstijl, werd gevraagd hardop na te denken tijdens het bestuderen van een tekst over sociale psychologie. Elk hardopdenkprotocol werd vervolgens verdeeld in zinvolle eenheden, waarna elke eenheid toegewezen werd aan een van de scoringscategorieën.

In overeenstemming met de theorie van Vermunt memoriseerden de reproductiegerichte studenten significant meer dan de betekenisgerichte en ongerichte studenten. Voor andere categorieën waren de verschillen tussen de leerstijlen minder duidelijk. Dit was waarschijnlijk een gevolg van de kleine steekproef en de gebruikte tekst.

1 Inleiding

In 1992 publiceerde Vermunt in zijn proefschrift de resultaten van zijn jarenlange onderzoek naar de manieren waarop studenten leren in het wetenschappelijk onderwijs, en hoe daarmee rekening zou kunnen worden gehouden bij de inrichting van het hoger onderwijs. Vermunt ontwikkelde in dat kader een leerstijldiagnose-instrument, de Inventaris Leerstijlen (ILS). Dit instrument is een vragenlijst met 120 items over studie-activiteiten, studiemotieven en studieopvattingen. Daarmee kan een leerstijlprofiel worden samengesteld. De diagnostische kant van het instrument be-

staat er uit dat studenten inzicht krijgen in de sterke en zwakke kanten van de eigen leerstijl.

Een leerstijl is volgens Vermunt (1992) een samenhangend geheel van leeractiviteiten die studenten gewoonlijk ontplooiën, de wijze waarop zij hun leerprocessen sturen, hun visie op leren en onderwijs en hun studie-motieven. Leerstijlen verschillen van elkaar op een vijftal aspecten: de wijze waarop studenten de studiestof cognitief verwerken, de wijze waarop deze verwerking wordt gereguleerd, de affectieve processen die zich bij het studeren voordoen, de mentale leermodellen die eraan ten grondslag liggen en de leeroriëntaties van de studenten. In dit onderzoek wordt alleen gekeken naar de cognitieve verwerkingsstrategieën en de regulatiestrategieën, de drie andere aspecten worden vooralsnog buiten beschouwing gelaten. De affectieve activiteiten worden in de ILS overigens buiten beschouwing gelaten. Tot de cognitieve verwerkingsstrategieën rekent Vermunt relateren en structureren, memoriseren, concrete verwerking, analyseren en kritisch verwerken. Als regulatiestrategieën beschouwt Vermunt de externe sturing van het leerproces en leerresultaten, de zelfsturing van het leerproces en leerresultaten, de zelfsturing van de leerinhoud en stuurloos leergedrag. Voor een uitgebreide, theoretische beschrijving van die aspecten en strategieën wordt verwezen naar Vermunt (1992).

Vermunt kwam tot vier, volgens hem wezenlijk verschillende leerstijlen. Ten eerste een *ongerichte* leerstijl: studenten met deze leerstijl zouden veel problemen hebben met het verwerken van de stof, geen grote hoeveelheden aankunnen en moeite hebben met het bepalen wat belangrijk is. Daarnaast zijn er studenten met een *reproductiegericht* leerstijl: deze studenten proberen de leerstof uit colleges en studieboeken zo goed als mogelijk te onthouden om die op tentamens te kunnen reproduceren. Voorts zijn er studenten wier studiegedrag zich kenmerkt door een *betekenisgericht*

leerstijl: zij pogen bedoelingen van auteurs te achterhalen, bredere verbanden te leggen, een eigen visie te ontwikkelen. Kortom, zij vertonen het meest een academische werkhouding. Ten vierde kunnen studenten onderscheiden worden die zich bedienen van een *toepassingsgerichte* leerstijl: deze studenten betrekken de studiestof op actuele verschijnselen, en gaan na wat ze ermee kunnen. In dit onderzoek is de toepassingsgerichte leerstijl achterwege gelaten, omdat deze leerstijl volgens het onderzoek van Vermunt (1992) en van Busato, Prins, Hamaker en Visser (1995) vrijwel niet direct in de verwerkings- en regulatiestrategieën zichtbaar wordt. Slechts de verwerkingsstrategie 'concreet verwerken' hangt enigszins samen met de toepassingsgerichte leerstijl. Daarnaast kan gezien de aard van de psychologie-opleiding worden verondersteld dat de toepassingsgerichte leerstijl meer in de eindfase van de studie voorkomt, hoewel empirische evidentie hiervoor ontbreekt. Met behulp van de ILS kunnen de vier leerstijlen in een puntenscore worden uitgedrukt. Hoewel een student derhalve van elke leerstijl kenmerken kan vertonen, veronderstelt Vermunt (1992, p. 72) dat één leerstijl domineert.

Het nadeel van zelfrapportage in het algemeen is dat sociale wenselijkheid als antwoordtendens de resultaten kan vertekenen (zie bijvoorbeeld Hoogstraten, 1979). Bovendien is het mogelijk dat studenten onvoldoende inzicht in hun eigen studiegedrag hebben. Men kan zich derhalve afvragen of de ILS misschien een leerstijl meet die men zich wenst of verbeeldt, en niet zozeer een leerstijl die men daadwerkelijk heeft. Doelstelling van dit onderzoek is een bijdrage te leveren tot de validatie van een gedeelte van de ILS: studeren studenten tijdens een reële studietaak op de manier waarop ze dat via de ILS beweren te doen? Om deze vraag te beantwoorden werd besloten om studenten een passage uit een studieboek, behorende tot het eerstejaars-curriculum, hardop denkend te laten bestuderen.

Tekstbestudering is een leertaak waarbij studenten grotendeels zelfstandig controle moeten uitoefenen over het verloop van hun leerproces (Vermunt, Lodewijks & Simons, 1986). Men kan bestuderingsprocessen trachten te meten kort na of tijdens de bestudering

van een tekst. Gegevens waarin ook de aard van de activiteit in termen van verschillende bestuderingsprocessen tot uiting komt, worden verkregen met introspectie of hardop-denken procedures. Het verschil tussen beide is dat de proefpersoon bij introspectie moet rapporteren over wat hij deed en dus wordt aangespoord te theoretiseren, terwijl bij hardop-denken procedures het theoretiseren zoveel mogelijk wordt tegengegaan (Wouters & De Jong, 1982).

Hoewel er veel kritiek is geleverd op het nut van verbale data en met name op de hardop-denken procedure (bijvoorbeeld Nisbett & Wilson, 1977), wordt de hardop-denken methode vaak gehanteerd in cognitief psychologisch onderzoek (onder andere Ericsson & Simon, 1980; Busato, Veenman & Elshout, 1993). Ericsson en Simon concluderen dat verbale rapportage de inhoud van cognitieve processen, die plaats vinden tijdens de taakuitvoering, met precisie weergeven. De onderzoeken die Nisbett en Wilson tot tegengestelde constatering brachten, hadden volgens Ericsson en Simon (1980) gemeen dat de proefpersonen ook het 'waarom' van hun processen moesten verbaliseren. Dit zou aanzetten tot theoretiseren en wordt daarom niet geschikt geacht om betrouwbare verbale data te verkrijgen (Ericsson & Simon, 1980).

Wouters en De Jong (1982) menen dat de hardop-denken methode gebruikt kan worden om meer directe informatie te verkrijgen over bestuderingsprocessen. Volgens hen moet er echter aan een aantal voorwaarden worden voldaan. Ten eerste moet er niet om specifieke informatie worden gevraagd waaraan de student anders geen aandacht had besteed. Dit om een sturende invloed van het hardop-denken te voorkomen. Verder dient de proefleider zich te beperken tot het stimuleren van het hardop-denken via niet-inhoudelijke opmerkingen als 'Wat denk je nu?' en 'Denk je eraan hardop te blijven denken?' Dit om het theoretiseren over cognitieve processen tegen te gaan.

In dit onderzoek wordt verwacht dat studenten met een betekenisgerichte leerstijl tijdens het hardop studeren bepaald studiegedrag vaker zullen laten zien dan studenten van de andere twee leerstijlgroepen, te weten: relateren, structureren, kritisch verwerken, concreet verwerken. Voorts wordt verwacht dat studenten

met een reproductiegerichte leerstijl meer zullen memoriseren en herhalen dan studenten met een niet-reproductie-gerichte leerstijl. Van studenten met een ongerichte leerstijl wordt verwacht dat zij meer stuurloos leergedrag zullen vertonen dan studenten met een niet-ongerichte leerstijl. Omdat Vermunt (1992, hoofdstuk 7) vond dat er een positieve samenhang bestaat tussen de betekenisgerichte leerstijl en tentamenresultaten, een negatieve samenhang tussen de ongerichte leerstijl en tentamenresultaten, en geen samenhang tussen de reproductiegerichte leerstijl en tentamenresultaten, wordt tenslotte verwacht dat de betekenisgerichte studenten na het bestuderen van een studietekst een hoger cijfer op een afsluitende toets zullen behalen dan de reproductiegerichte en ongerichte studenten.

2 Methode

2.1 Proefpersonen

Proefpersonen waren ruim driehonderd eerstejaars psychologie-studenten van de Universiteit van Amsterdam. De vaststelling van de leerstijl geschiedde met behulp van de ILS. Tijdens verplichte werkgroepen, behorend bij het vak Inleiding in de Psychologie, werd de ILS afgenomen. Deze meting vond plaats in oktober, vlak na het eerste tentamen.

In een replicatie van het promotie-onderzoek van Vermunt (1992) aan de Universiteit van Amsterdam (Busato et al., 1995) werd vastgesteld dat de strategieën relateren en structureren, concrete verwerking, kritische verwerking en zelfsturing leerproces hoog laadden op de leerstijl *betekenisgericht*. De strategieën memoriseren en herhalen, externe sturing leerproces en externe sturing leerresultaten laadden hoog op de leerstijl *reproductiegericht*. De strategie stuurloos leergedrag laadde hoog op de leerstijl *ongericht*, evenals een mentaal leermodel stimulerend onderwijs en samen studeren. Busato et al. (1995) vonden vrijwel identieke factorladingen als Vermunt (1992). Als operationalisatie van de drie leerstijlen werd daarom gekozen voor de opgetelde scores van de genoemde schalen die de hoogste lading hadden op de desbetreffende leerstijl. Van alle studenten werden de scores op de drie leerstijlen berekend.

De ILS levert per student een score op elke leerstijl. Dat betekent dat het mogelijk is dat een student hoog scoort op meer dan één leerstijl. Het is zelfs mogelijk dat een student hoge scores haalt op elk van de vier leerstijlen van de ILS. Voor het onderzoek wilden we gebruik maken van proefpersonen die hoog scoren op één van de drie te valideren leerstijlen, en laag op de beide andere. Hierbij werd als pragmatisch criterium gehanteerd dat de student minimaal een halve standaardafwijking boven het gemiddelde moest scoren op de desbetreffende leerstijl en onder het gemiddelde op de andere twee leerstijlen. Op de betekenisgerichte schaal (36 items, maximale score van 180) werd voor de hele groep eerstejaars studenten een gemiddelde score gevonden van 110.30, met een standaardafwijking van 20.04 ($N=320$). Op de reproductiegerichte schaal (16 items, maximale score van 80) werd een gemiddelde waarde gevonden van 44.27 met een standaardafwijking van 8.64 ($N=327$). Op de schaal ongericht (27 items, maximale score van 135) was het gemiddelde 67.35 met een standaardafwijking van 13.95 ($N=321$). Aan de hand van de scores op deze afname werden 82 studenten aangeschreven met een verzoek tot deelname aan het experiment getiteld 'Tekstlezen'.

2.2 Materiaal en procedure

In dit experiment kreeg de proefpersoon een tekst van tien pagina's voorgelegd uit het boek 'Sociale Psychologie' van Meertens en Von Grumbkow (1992). Dit hoofdstuk handelde over het onderwerp 'macht'. In de instructie werd de proefpersoon gevraagd de tekst te bestuderen zoals dat in een reguliere studiesituatie zou worden gedaan. Tevens werd verzocht bij het studeren hardop na te denken. Dit hield in dat alles waar tijdens het bestuderen van de tekst aan werd gedacht hardop diende te worden verwoord. Alles wat de proefpersoon hierbij uitsprak werd opgenomen met behulp van een cassette recorder. Om de proefpersoon eraan te herinneren hardop na te blijven denken stonden hier en daar in de tekst zwarte blokjes. Deze waren bedoeld als geheugensteuntje en fungeerden niet als directe opdracht tot hardop denken. Indien de proefpersoon tijdens het bestuderen van de tekst langer dan twee minuten stil was, herinnerde de proefleider de proefper-

soon eraan hardop na te blijven denken. Nadat de proefpersoon de tekst had bestudeerd, kreeg deze een tentamen voorgelegd dat bestond uit vijf open en vijf multiple choice vragen. Dit was in de instructie reeds aan hen medegedeeld. De proefpersonen mochten zo lang studeren en zo lang over het tentamen doen als ze zelf wilden.

De proefpersonen werd niet meegedeeld dat de helft van de vragen vooral gericht was op kennis en de andere helft vooral op begrip. Met de kennisvragen werd naar een in de tekst vermeld feit gevraagd. Voor het beantwoorden van een begripsvraag moest feitenkennis worden toegepast of moesten er verbanden worden gelegd. Voor iedere vraag kon de proefpersoon maximaal één punt behalen. De eindscore werd berekend door het optellen van de scores op de verschillende vragen. Een voorbeeld van een meerkeuze vraag, gericht op kennis, is: "Welk machtsmiddel komt *niet* voor in de indeling van Foa en Foa?" (alternatieven a) hoop, b) liefde, c) status en d) onzekerheidsreductie). Een voorbeeld van een open vraag, gericht op begrip, luidt: "Wat wordt bedoeld met de machtsafstandreductietendens en wanneer neemt deze volgens Mulder toe?"

Het in dit onderzoek gebruikte design is het Static-Group-Comparison design (Judd, Smith & Kidder, 1991).

2.3 Dataverwerking

In de bestudeerde literatuur over hardop-denken-protocollen werd in het artikel van Wouters en De Jong (1982) een schema met scoringscategorieën gevonden dat betrekking had op soortgelijke strategieën die in dit onderzoek zouden worden gemeten. Ook Vermunt (1986) bouwt voort op Wouters en De Jong. De categorieën van Wouters en De Jong bleken zeer sterk overeen te komen met de vragen van Vermunt in de ILS. De hoofdcategorieën in dit onderzoek waren relateren en structureren, concreet verwerken, kritisch verwerken, zelfsturing, memoriseren en onbegrip constateren (stuurloos leergedrag). De score voor elke hoofdcategorie was de optelling van de specifiekere subcategorieën (zie bijlage A). Binnen de scoringscategorie 'zelfsturing leerprocessen' werd er nog een aparte categorie 'herformuleren' opgesteld. Deze werd berekend door het aantal malen dat 'woord herformuleren', 'zin herformu-

leren', 'passage herformuleren' en 'beeldmateriaal herformuleren' werd geturfd, bij elkaar op te tellen.

Tevens werd een aantal variabelen toegevoegd aan de scoringscategorieën. Deze hadden betrekking op uitspraken die niet binnen het schema van Wouters en De Jong pasten en niet binnen de bovenstaande hoofdcategorieën onder te brengen waren. Ze refereerden aan activiteiten waarvan verwacht werd dat ze met een bepaalde leerstijl zouden samenhangen. Deze variabelen zijn 'constateren eigen handelingen', 'begrip constateren', 'herlezen', 'werkwijze', 'markeringen in en aantekeningen naast de tekst' en een 'restcategorie', waarin de activiteiten gescoord werden die niet in één van de overige scoringscategorieën onder te brengen waren. In bijlage A worden de verschillende scoringscategorieën nader omschreven, alsmede relevante voorbeelden gegeven.

Alle protocollen werden uiteindelijk door de eerste twee auteurs gescoord. Voorafgaande aan het daadwerkelijke scoren van de protocollen werden twee trainingsdagen belegd om het scoren te oefenen. De trainingen bestonden uit het scoren van eenzelfde pilot-protocol, waarna de individuele scoringen werden besproken. De discussies hadden als doel het verkrijgen van overeenstemming tussen de beoordelaars over de wijze van scoren. De scoring van deze categorieën vond plaats door ieder uitgetypt hardop-denken-protocol in zinvolle scoringsenheden te verdelen en vervolgens te bepalen welke van de scoringscategorieën op die eenheid van toepassing was. Voor de indeling van zinvolle eenheden werd aangesloten bij Wouters en De Jong (1982). Binnen elke zinvolle eenheid vond één studieactiviteit plaats. Het probleem van de afbakening is derhalve tegelijkertijd met de benoeming van de activiteit, de indeling in de scoringscategorie, opgelost. Steeds stelde de beoordelaar zichzelf de vraag welke bewerking de student toepaste op de aangeboden informatie. Een zinvolle eenheid kon bestaan uit een woord, een zin, zinsdeel of een passage. Ter bepaling van de interbeoordelaars-betrouwbaarheid werden 9 van de 39 protocollen tweemaal gescoord. Hierbij werd door de tweede beoordelaar de indeling in eenheden aangehouden die de eerste beoordelaar had opgesteld. De interbeoorde-

laars-betrouwbaarheid werd uitgedrukt als κ (kappa), een coëfficiënt voor nominale schalen. Deze kan direct geïnterpreteerd worden als het gedeelte van de scores die met elkaar overeenkomen, na correctie voor toevalsscores. De waarde van κ varieert tussen 0 en 1 (Cohen, 1960). Er is gekozen om niet per subcategorie de interbeoordelaars-betrouwbaarheid te berekenen, maar per totale categorie. Concreet betekende dit dat wanneer de eerste beoordelaar bijvoorbeeld een eenheid benoemde als een MAB (zie bijlage A), en de tweede beoordelaar deze eenheid scoorde als MCM, er toch sprake was van een perfecte overeenstemming omdat beide subcategorieën behoren tot de overkoepelende categorie memoriseren.

Tenslotte werd voor elke hoofdcategorie en de bijbehorende subcategorieën gekeken of de gemiddelden van de drie leerstijlgroepen van elkaar verschilden. Dit werd gedaan voor zowel de absolute scores als de relatieve scores. Het relatieve aantal eenheden werd berekend door voor elke proefpersoon per (sub)categorie het bijbehorende aantal eenheden te delen door het totaal aantal eenheden van de desbetreffende proefpersoon. Het relatieve aantal eenheden is interessant, omdat dan kan worden gekeken hoe vaak een bepaald studiegedrag voorkomt in verhouding tot het totaal aantal gedragingen.

De gegevens werden geanalyseerd met SPSS en met Statview, een statistisch pakket voor Apple Macintosh-computers.

3 Resultaten

3.1 Steekproef

Er reageerden 31 studenten, twee daarvan kwamen niet opdagen en twee anderen bleken niet in staat te zijn om hardop na te denken. De verdeling over de leerstijlen was als volgt: acht studenten werden als betekenisgericht aangemerkt, vijf als reproductiegericht en veertien als ongericht. In Tabel 1 staan de gemiddelde leerstijlscores per groep.

De gemiddelde leeftijd van de proefpersonen was 21.7 jaar, met een standaarddeviatie van 3.25. Een t-toets wees uit dat de ongerichte studenten significant jonger waren dan de betekenisgerichte en reproductiegerichte studenten ($Xb + r = 22.8$, $sd = 4.04$; $Xo = 20.7$, $sd = 1.94$; $t(16.97) = 1.73$; $p = .05$). Er deden twee keer zoveel vrouwen als mannen mee (respectievelijk 18 om 9).

3.2 Leertijd

De studenten bestudeerden de tekst gemiddeld 44.3 minuten, met een standaarddeviatie van 14.1. De gemiddelde leertijd van de drie groepen verschildte niet van elkaar.

3.3 De protocolanalyses

Om de protocolanalyses zinvol te interpreteren werd allereerst gekeken naar de interbeoordelaarsbetrouwbaarheid. De Cohens kappa's van de negen dubbelgescoorde protocollen lagen tussen .57 en .90. De gemiddelde waarde bedroeg .77. Deze waarde wordt als voldoende hoog geacht om te veronderstellen dat beide beoordelaars de resterende protocollen op dezelfde wijze gescoord hebben.

Het gemiddelde aantal eenheden voor de

Tabel 1
Gemiddelden en standaarddeviaties van de leerstijlscores per groep

Leerstijl	Betek-score	Repro-score	Ongericht-score
Betekenisgericht (N=8)	140.63 (8.48)	37.88 (5.79)	53.86 (13.46)
Reproductiegericht (N=5)	98.00 (3.24)	56.40 (2.30)	58.00 (11.79)
Ongericht (N=14)	94.86 (10.52)	40.50 (3.65)	82.71 (6.22)
Hele populatie	110.30 (20.04) (N=320)	44.27 (8.64) (N=327)	67.35 (13.95) (N=321)

Tabel 2

Gemiddelden en standaarddeviatie per hoofdcategorie voor de drie groepen. Op de eerste regel staat het absolute aantal, daaronder het relatieve

Leerstijl	Betekenisgericht	Reproductiegericht	Ongericht
Structureren (niet inhoud.)	14.38 (13.97) .082 (.049)	12.40 (6.58) .075 (.036)	8.43 (8.56) .084 (.068)
Relateren (inhoudelijk)	38.00 (47.19) .229 (.176)	31.00 (9.19) .188 (.044)	20.36 (13.70) .182 (.087)
Structureren/relateren	52.38 (56.31) .31 (.176)	43.40 (10.64) .263 (.04)	28.57 (18.66) .264 (.088)
Concreet verwerken	16.50 (12.90) .117 (.069)	8.40 (7.54) .048 (.045)	9.71 (10.16) .09 (.059)
Kritisch verwerken	5.50 (5.03) .048 (.068)	3.60 (2.61) .024 (.019)	4.86 (6.11) .045 (.039)
Herformuleren	29.63 (28.61) .208 (.161)	36.00 (27.23) .196 (.109)	29.21 (23.50) .29 (.194)
Zelfsturing	45.75 (28.18) .334 (.112)	46.60 (32.11) .249 (.117)	36.79 (22.55) .074 (.115)
Memoriseren	18.38 (19.57) .131 (.142)	39.80 (17.98) .249 (.117)	12.86 (22.80) .074 (.115)
Onbegrip constateren	6.38 (4.81) .054 (.043)	5.2 (4.60) .03 (.024)	1.93 (1.94) .029 (.038)
Totaal	149.63 (116.91)	167.60 (45.71)	106.43 (65.12)

gehele steekproef bedroeg 130.56, met een standaarddeviatie van 82.53, een minimum van 24 eenheden en een maximum van 403. Vervolgens werd gekeken of de drie leerstijlgroepen verschilden in het totaal aantal ingedeelde eenheden. Een variantieanalyse wees uit dat het niet het geval was. Wel bleek dat het verschil tussen de studenten met een ongerichte leerstijl en de andere twee leerstijlgroepen significantie benaderde. Bij ongerichte studenten werden minder eenheden ingedeeld ($X_o=106.43$, $sd=65.12$; $X_b+r=156.54$, $sd=93.55$; $t(25)=1.63$; $p=.058$). Omdat dit verschil dusdanig dicht tegen de kritieke grens van $\alpha=.05$ ligt, is besloten niet alleen te kijken naar de verschillen voor het absoluut aantal eenheden, maar ook naar de verschillen voor het relatieve aantal. In Tabel 2 zijn de gemiddelden en standaarddeviaties van het totaal aantal eenheden per groep weergegeven. In deze tabel staan ook per groep de gemiddelden en standaarddeviaties per hoofdcategorie weergegeven, zowel absoluut als relatief.

Hieronder zullen de analyses per hoofdcate-

gorie worden besproken. Eventuele significante verschillen voor de subcategorieën worden ook vermeld. Wanneer voor een variable bij het relatieve aantal eenheden een significant verschil werd gevonden, en voor het absolute aantal eenheden het verschil in de juiste richting wees, worden voor de volledigheid ook de toetsingsgegevens voor het absolute aantal eenheden gerapporteerd, en vice versa. Naast variantie-analyses zullen ook eenzijdige t-toetsen gerapporteerd worden, omdat op grond van de theorie van Vermunt (1992) voor elke studeeractiviteit verwacht wordt dat deze kenmerkend is voor één leerstijlgroep. Deze leerstijlgroep zal de desbetreffende studeeractiviteit volgens verwachting vaker vertonen dan beide andere leerstijlgroepen. Omdat de groepen verschillen in grootte, kan het bij de t-toetsen voorkomen dat de varianties binnen de groepen ongelijk zijn. Wanneer dat het geval is, moet worden gekeken naar de 'separate variance estimate', die SPSS naast de 'pooled variance estimate' levert (zie een willekeurige SPSS-manual). Bij de t-waarde wordt dan een

gebroken aantal vrijheidsgraden gerapporteerd.

3.4 Structureren en relateren

Een variantie-analyse liet geen significant verschil zien tussen de drie leerstijlgroepen voor de categorieën 'structureren' en 'relateren'. Dit gold voor zowel het absolute aantal eenheden als voor het relatieve aantal.

Tegen de verwachting in toonde ook een t-toets geen verschil aan tussen betekenisgerichte studenten enerzijds en reproductiegerichte en ongerichte studenten anderzijds voor de hoofdcategorieën 'structureren' en 'relateren'.

Voor de subcategorie 'overkoepelen' werd een onverwacht resultaat gevonden. Een variantie-analyse wees uit dat er een significant verschil bestond tussen de drie groepen voor de absolute score ($X_r=2.8$, $sd=1.79$; $X_o=.43$, $sd=.85$; $X_b=1.63$, $sd=1.85$; $F(2,24)=5.87$; $p<.01$). Een post-hoc Newman-Keuls procedure liet zien dat alleen het verschil tussen de reproductiegerichte en de ongerichte studenten significant was. De betekenisgerichte studenten verschilden niet van de beide andere groepen. Voor de relatieve score werd dit onverwachte verschil niet gevonden.

3.5 Concreet verwerken

Ook bij de hoofdcategorie 'concreet verwerken' liet een variantie-analyse geen significante verschillen tussen de groepen zien, voor zowel het absolute als het relatieve aantal eenheden.

3.6 Kritisch verwerken

Tegen de verwachting in werden voor de hoofdcategorie 'kritisch verwerken' en de afzonderlijk subcategorieën geen significante verschillen gevonden tussen de drie groepen, niet voor het absolute aantal eenheden en niet voor het relatieve. Ook de t-toetsen lieten geen significante verschillen zien tussen de betekenisgerichte groep enerzijds en de reproductiegerichte groep en ongerichte groep anderzijds.

3.7 Herformuleren en zelfsturing

Een variantie-analyse wees uit dat er voor de hoofdcategorieën 'herformuleren' en 'zelfsturing' geen significant verschil werd gevonden, niet voor het absolute aantal eenheden en niet voor het relatieve.

Verwacht werd dat de betekenisgerichte studenten vaker zouden plannen dan de andere twee groepen studenten. Uit een t-toets bleek echter dat voor de subcategorie 'plannen' ongerichte studenten absoluut gezien significant minder vaak planden dan de andere twee leerstijlgroepen ($X_o=4.57$, $sd=4.77$; $X_b+r=9.69$, $sd=8.77$; $t(18.22)=1.87$; $p<.05$). Dat verschil werd niet gevonden voor het relatieve aantal eenheden.

3.8 Memoriseren

Een variantie-analyse wees uit dat er een verschil bestond voor de hoofdcategorie 'memoriseren' voor het relatieve aantal eenheden ($X_r=.249$, $sd=.117$; $X_b=.131$, $sd=.142$; $X_o=.074$, $sd=.115$; $F(2,24)=3.69$, $p<.05$). Een post-hoc Newman-Keuls procedure liet zien dat het verschil tussen de reproductiegerichte en de ongerichte studenten significant was. De betekenisgerichte studenten verschilden niet van de beide andere leerstijlgroepen. Een variantie-analyse voor de absolute score wees op een verschil in de verwachte richting ($X_r=39.80$, $sd=17.98$; $X_b=18.38$, $sd=19.57$; $X_o=12.86$, $sd=22.80$; $F(2,24)=3.01$, $p=.068$).

Uit de t-toetsen bleek dat, volgens verwachting, reproductiegerichte studenten significant meer memoriseren dan de studenten met een andere leerstijl (absoluut: $X_r=39.8$, $sd=17.98$; $X_b+o=14.68$, $sd=21.38$; $t(25)=-2.41$; $p<.05$; relatief: $X_r=.247$, $sd=.117$; $X_b+o=.095$, $sd=.125$; $t(25)=-2.50$; $p<.05$). Ook voor de subcategorieën 'abstracte begrippen herlezen' (absoluut: $X_r=28.6$, $sd=10.67$; $X_b+o=11.18$, $sd=16.66$; $t(25)=-2.22$; $p<.05$; relatief: $X_r=.176$, $sd=.064$; $X_b+o=.069$, $sd=.097$; $t(25)=-2.32$; $p<.05$) en de subcategorie 'concreet materiaal herlezen' (absoluut: $X_r=11.2$, $sd=10.55$; $X_b+o=3.41$, $sd=6.86$; $t(25)=2.08$; $p<.05$; relatief: $X_r=.073$, $sd=.072$; $X_b+o=.024$, $sd=.043$; $t(25)=2.01$; $p<.05$) werd dit verschil gevonden.

Tegen de verwachting in bleek dat voor de subcategorie 'abstracte begrippen herhalen' ongerichte studenten absoluut gezien significant minder vaak abstracte begrippen herhaalden dan de andere twee leerstijlgroepen ($X_o=8.57$, $sd=16.48$; $X_b+r=20.69$, $sd=15.80$; $t(25)=1.95$; $p<.05$). Dit verschil werd ook voor het relatieve aantal gevonden ($X_o=.047$, $sd=.074$; $X_b+r=.135$, $sd=.106$; $t(25)=2.51$; $p<.05$).

3.9 Onbegrip constateren

Uit de variantieanalyse bleek dat er voor het absolute aantal eenheden een verschil bestond tussen de drie leerstijlgroepen voor de hoofdcategorie 'onbegrip constateren' ($X_b=6.38$, $sd=4.81$; $X_r=5.2$, $sd=4.60$; $X_o=1.93$, $sd=1.94$; $F(2,24)=4.54$; $p<.05$). Een post-hoc Newman-Keuls procedure liet zien dat de betekenisgerichte studenten meer onbegrip constateerden dan de ongerichte. De reproductiegerichte studenten verschilden niet van de beide andere leerstijlgroepen. Ook voor de subcategorie 'uitleg aan de proefleider vragen' is een voor het absolute aantal eenheden een verschil tussen de leerstijlgroepen gevonden ($X_b=3.38$, $sd=4.07$; $X_r=2.6$, $sd=3.13$; $X_o=.43$, $sd=.94$; $F(2,24)=3.54$; $p<.05$). Een post-hoc Newman-Keuls procedure liet zien dat ook hier het verschil tussen de betekenisgerichte en de ongerichte studenten significant was. De reproductiegerichte studenten verschilden niet van de beide andere leerstijlgroepen. Deze twee bovenstaande verschillen waren tegen verwachting.

3.10 Toegevoegde variabelen

Voor twee van de zes toegevoegde variabelen zijn significante verschillen tussen de leerstijlgroepen gevonden. Een variantieanalyse wees uit dat er voor het absolute aantal eenheden een verschil bestond tussen de drie leerstijlgroepen voor de variabele 'werkwijze' ($X_b=2.00$, $sd=1.77$; $X_r=1.20$, $sd=1.30$; $X_o=.29$, $sd=.61$; $F(2,24)=5.48$; $p<.05$). Een post-hoc Newman-Keuls procedure liet zien dat het verschil tussen de betekenisgerichte studenten en de ongerichte studenten significant was. Betekenisgerichte studenten verwezen meer naar hun werkwijze dan de ongerichte. De reproductiegerichte studenten verschilden niet van de beide andere leerstijlgroepen. Voor het rela-

tieve aantal eenheden wees het verschil in dezelfde richting ($X_b=.020$, $sd=.020$; $X_r=.009$, $sd=.013$; $X_o=.004$, $sd=.010$; $F(2,24)=3.17$; $p=.060$). Ook hier liet een post-hoc Newman-Keuls procedure zien dat het verschil tussen de betekenisgerichte studenten en de ongerichte studenten significant was.

De ongerichte studenten scoorden wat betreft het absolute aantal eenheden significant lager op de variabele 'begrip constateren' dan de andere twee leerstijlgroepen ($X_o=2.21$, $sd=3.87$; $X_b+r=6.62$, $sd=7.28$; $t(17.98)=1.94$; $p<.05$). Voor het relatieve aantal wees het verschil in dezelfde richting ($X_o=.018$, $sd=.029$; $X_b+r=.033$, $sd=.026$; $t(25)=1.42$; $p=.084$).

3.11 Tentamen

De tentamens werden door twee beoordelaars gescoord. De correlatie tussen de scores van de multiple choice vragen was perfect, de correlatie tussen de scores van de open vragen bedroeg $r=.98$. Een variantieanalyse wees uit dat de gemiddelde tentamencijfers (meerkeuze, open vragen en totaal) voor betekenisgerichte, reproductiegerichte en ongerichte studenten niet significant verschilden. In Tabel 3 staan de gemiddelde tentamencijfers per groep weergegeven.

4 Discussie

Doel van dit onderzoek was om een bijdrage te leveren aan de validatie van delen van de Inventaris Leerstijlen (ILS) van Vermunt. De verschillen in verwerkingsstrategieën en regulatiestrategieën die de ILS beoogt te meten, zouden tijdens de studietoek waarbij studenten gevraagd werd hardop te denken, waarneembaar moeten zijn. De resultaten van dit onderzoek wijzen slechts gedeeltelijk in die richting.

Tabel 3
Gemiddelden en standaarddeviaties van de tentamencijfers per groep

Leerstijl	Meerkeuze	Open vragen	Totaal
Betekenisgericht	3.00 (1.20)	3.16 (1.14)	6.16 (2.03)
Reproductiegericht	3.40 (1.14)	2.80 (0.21)	6.20 (1.26)
Ongericht	2.86 (1.41)	2.39 (1.17)	5.24 (2.18)
Hele steekproef	3.00 (1.27)	2.69 (1.08)	5.69 (1.99)

Voor de verwerkingsstrategie 'memoriseren' bleek dat voor zowel de totale categorie 'memoriseren' als de twee belangrijkste subcategorieën daarvan, 'abstract begrippen herlezen' en 'concreet materiaal herlezen', de verschillen significant en in de gevonden richting waren. Voor de hoofdcategorie 'onbegrip constateren' zijn de resultaten tegengesteld aan wat verwacht werd.

Er kan een aantal redenen worden aangevoerd waarom significante verschillen niet voor alle categorieën optraden. Allereerst kan de manier van operationaliseren van de subschalen van de ILS genoemd worden. Het operationaliseren van 'stuurloos leergedrag' bijvoorbeeld, een kenmerk van studenten met een ongerichte leerstijl, bleek moeilijk te zijn. Hoe kan worden gemeten of studenten *niet* sturen? Is dat te turven? In dit onderzoek is daarom voor deze subschaal als operationalisatie 'onbegrip constateren' gekozen. Uit de resultaten bleek echter dat juist studenten met een betekenisgerichte leerstijl significant meer onbegrip constateerden dan de studenten met een ongerichte leerstijl. 'Onbegrip constateren' kan men, achteraf beschouwd, zien als een regulatie-activiteit. Men kan zich voorstellen dat het constateren van onbegrip aanzet om zichzelf te sturen tot het bijvoorbeeld nogmaals lezen van een tekstdeel om meer begrip te verkrijgen. Met andere woorden, onbegrip constateren kan aanleiding zijn om te plannen, en dat is een kenmerk van studenten met een betekenisgerichte leerstijl.

Een andere reden voor het niet optreden van significante verschillen op bepaalde categorieën kan zijn dat de gebruikte studietekst zich niet leent voor dat betreffende leergedrag. Men kan daarbij vooral denken aan de categorie 'kritisch verwerken'. Er werd door de studenten tijdens de studietaak relatief weinig kritisch verwerkt, minder dan welk ander studiegedrag. Kritisch verwerken heeft in de vragenlijst van Vermunt onder andere betrekking op het kritisch bekijken van interpretaties van deskundigen, en het nagaan of de conclusies van de auteurs van een studieboek logisch volgen uit de feiten. In de voor dit onderzoek gebruikte studietekst werden enkele theorieën over macht vrij summier beschreven. Voor een kritische blik heeft men waarschijnlijk meer gedetailleerde informatie nodig. De auteur van de

studietekst koos bovendien geen partij voor één van de theorieën, maar gaf ze achter elkaar weer. Dan wordt het moeilijk om de conclusies van de auteurs kritisch te bekijken. Verder lijkt de setting van het onderzoek voor 'kritisch verwerken' van belang. Men kan zich voorstellen dat een eerstejaars student moeite heeft met het hardop kritisch verwerken als er een ('wijzere') ouderejaars student als proefleider bijzit. Aan de andere kant zou 'kritisch verwerken' echter één van de leergedragingen zijn die betekenisgerichte studenten bij zichzelf tijdens de zelfreportage in de ILS overschatten. Ook is het mogelijk dat pas in een later stadium van de studie daadwerkelijk verschillen tussen de leerstijlgroepen te constateren zijn in regulier studeergedrag voor wat betreft 'kritisch verwerken'.

Volgens Vermunt (1992) is het onderscheid tussen de leerstijlgroepen beter waarneembaar als het om grote hoeveelheden studietekst gaat. De gebruikte tekst telde tien pagina's, een relatief kleine hoeveelheid. Dat kan een reden zijn waarom de verschillen tussen de leerstijlgroepen bij sommige categorieën, zoals 'structureren', niet waarneembaar zijn. Bovendien kunnen studenten afhankelijk van deze studietaak voor een andere strategie kiezen. Tien pagina's zijn relatief makkelijk uit het hoofd te leren, structuur aanbrengen en selecteren wat belangrijk is, is dan minder nodig. Desondanks werden voor vrijwel alle categorieën (behalve 'kritisch verwerken' en 'onbegrip constateren') verschillen gevonden die, hoewel niet significant, wel wezen in de verwachte richting, voor de totale categorie en/of één van de subcategorieën. Deze bevindingen zijn aanleiding te veronderstellen dat in toekomstig onderzoek met grotere teksten de verschillen duidelijker zullen blijken.

Opvallend was de moeite die sommige studenten hadden met het hardop studeren. De kwaliteit van de hardop-denken-protocollen liep nogal uiteen, wat onder meer te zien is aan de hoge standaarddeviatie bij het totaal aantal eenheden en het grote verschil tussen het minimum en het maximum. Sommige studenten bleven tijdens het hardop studeren vrijwel aan één stuk praten terwijl andere studenten lange tijd hun mond hielden en aangespoord moesten worden. Uit de literatuur zijn dergelijke problemen bekend met hardop-denken-protocollen (zie bijvoorbeeld Ericsson & Simon, 1980). Boven-

dien is tussen de leerstijlgroepen een verschil gevonden dat significantie benaderde voor het totaal aantal zinvolle eenheden waarin de protocollen werden ingedeeld: bij ongerichte studenten werden minder eenheden ingedeeld dan bij de andere twee groepen studenten. Dat is evenwel te verwachten, omdat voor bijna alle subcategorieën voorspeld werd dat betekenisgerichte of reproductiegerichte studenten het betreffende gedrag méér zouden laten zien dan de andere twee groepen. Desalniettemin, het grote verschil tussen studenten in het absolute aantal eenheden pleit ervoor ook waarde te hechten aan de verschillen tussen de leerstijlgroepen voor wat betreft het relatieve aantal. Reproductiegerichte studenten bijvoorbeeld memoriseerden absoluut gezien meer dan de andere twee groepen. Dit zou het gevolg kunnen zijn van het feit dat bij deze groep totaal gezien meer eenheden zijn onderscheiden. Wellicht treedt het verschil niet op als gekeken wordt naar het relatieve aantal. Uit de resultaten blijkt echter dat het verschil ook voor het relatieve aantal eenheden bij memoriseren optreedt. Bij 'plannen' en 'zichzelf een vraag stellen' daarentegen is dat niet het geval. Studenten met een ongerichte leerstijl plannen significant minder vaak dan de andere twee groepen. Relatief gezien plannen ze echter even veel. Men zou kunnen concluderen dat ongerichte studenten zelfsturend studiegedrag absoluut gezien minder vertonen, maar dat het verschil verhoudingsgewijs meevalt.

Ten slotte kan de relatief kleine steekproef genoemd worden als reden voor het niet optreden van significante verschillen op bepaalde categorieën. Het is bekend dat hoe kleiner de groepen zijn die men met elkaar vergelijkt, hoe kleiner het onderscheidingsvermogen is. We pleiten daarom voor een replicatie van dit onderzoek met meer proefpersonen. In dit onderzoek is de steekproef klein uitgevallen omdat onder de eerstejaars studenten er maar weinig waren die (extreem) hoog scoorden op slechts één leerstijl en dus weinig die voor het onderzoek uitgenodigd konden worden.

Het feit dat ongerichte studenten significant jonger waren dan de andere twee groepen is in overeenstemming met de theorie van Vermunt (1992). Volgens de resultaten uit zijn onderzoek vertonen studenten naarmate ze ouder worden meer kenmerken van de betekenisge-

richte leerstijl. Longitudinaal onderzoek zou hier meer over kunnen uitwijzen.

Er zijn in dit onderzoek tevens significante verschillen gevonden die niet voorspeld zijn. Bij de subcategorie 'overkoepelen' vertoonden reproductiegerichte studenten dit gedrag significant meer dan de ongerichte. De reden voor het optreden van dit verschil blijft onduidelijk. Verder bleek dat bij de subcategorie 'plannen' de ongerichte studenten dit leergedrag significant minder vertoonden dan de andere twee groepen, terwijl op grond van het proefschrift van Vermunt (1992) werd voorspeld dat betekenisgerichte studenten meer zouden plannen dan de studenten met een andere leerstijl. Volgens Vermunt hangen de meeste subschalen van zijn vragenlijst samen met één bepaalde leerstijl. Een verschil tussen de andere leerstijlen (hier reproductiegerichte en ongerichte studenten) zou er dus niet zijn. Desondanks is voor 'plannen' zo'n verschil in ons onderzoek wel gevonden. Ook is iets dergelijks het geval voor de subcategorie 'abstract materiaal herlezen'. Ongerichte studenten doen dit minder vaak dan de andere twee groepen, terwijl verwacht werd dat reproductiegerichte studenten juist vaker abstract materiaal zouden herhalen. Een verschil tussen betekenisgerichte en ongerichte studenten op deze studieactiviteit is onverwacht op grond van Vermunts theorie.

Wat het tentamen betreft werden geen significante verschillen gevonden. Wel wezen bij de open vragen de verschillen tussen de leerstijlgroepen in de verwachte richting. De betekenisgerichte studenten scoorden het hoogst en de ongerichte studenten het laagst. Dat kan als een mogelijke ondersteuning van de selectie van de leerstijlgroepen worden gezien.

Dit onderzoek kan gezien worden als een eerste stap van het valideren van de volledige ILS. Het valideren van deze vragenlijst is nodig omdat te verwachten valt dat de ILS op grote schaal gebruikt gaat worden als diagnostisch instrument. De Landelijke Studievoordigheidsdag in Groningen (april 1994) was daar een illustratie van.

De hardop-denken-procedure lijkt voor verder validatie-onderzoek aan te raden. Te denken valt aan onderzoek per specifiek studeergedrag, zoals kritisch verwerken. Er zouden bij-

voorbeeld drie teksten genomen kunnen worden, die uitnodigen tot het innemen van een standpunt. Voor de strategie structureren kan gedacht worden aan een veel grotere tekst, waarin een lezer zelf duidelijk structuur moet aanbrengen. Het is dan te verwachten dat betekeniserichte studenten dit anders zullen doen dan de andere twee groepen.

Over de studenten met een ongerichte leerstijl is weinig bekend. Men weet vooral wat ze niet doen. Daarom kan voor vervolgstudies gedacht worden aan een onderzoek waarin de mentale leermodellen en leeroriëntaties geoperationaliseerd worden, hoe moeilijk dat ook lijkt, om zodoende een beter beeld te kunnen krijgen van het studeergedrag van de groep studenten met die ongerichte leerstijl. Zowel Vermunt (1992) als Busato et al. (1995) vonden hoge factorscores voor de ongerichte leerstijl op de mentale leermodellen en leeroriëntaties. Uit de factoranalyses kwam eveneens naar voren dat de ongerichte leerstijl niet samenhangt met één van de verwerkingstrategieën. Toch zullen deze mensen strategieën hanteren. Het zou interessant zijn er achter te komen welke. Door onderzoek te doen bij studenten met een ongerichte leerstijl kan men wat betreft de verwerkings- en regulatiestrategieën meer te weten komen over die studenten. Nu kan alleen geconstateerd worden dat de ongerichte leerstijl samenhangt met de weinig inzichtelijke regulatiestrategie 'stuurloos leergedrag'.

Literatuur

- Busato, V.V., Prins, F.J., Hamaker, C., & Visser, K. (1995). Leerstijlonderzoek gerepliceerd; de samenhang tussen leerstijlen en intelligentie. *Tijdschrift voor Onderwijsresearch*, 20, 332-340.
- Busato, V.V., Veenman, M.V.J., & Elshout, J.J. (1993). Metacognitieve instructies in een computer-ondersteunde leeromgeving. *De Psycholoog*, 28, 1, 13-17.
- Cohen, J.A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Hoogstraten, J. (1979). *De machteloze onderzoeker: voetangels en klemmen van sociaal-wetenschappelijk onderzoek*. Meppel/Amsterdam: Boom.

- Judd, C.M., Smith, E.R., & Kidder, L.H. (1991). *Research methods in social relations (sixth edition)*. Fort Worth, TX: Holt, Rinehart and Winston.
- Meertens, R.W., & Grumbkow, J. von (Red.) (1992). *Sociale Psychologie*. Groningen: Wolters-Noordhoff.
- Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Vermunt, J.D.H.M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs*. Amsterdam/Lisse: Swets & Zeitlinger.
- Vermunt, J.D.H.M. (1986). Hardop-denken als onderzoeksmethode naar regulatieprocessen bij tekstbestudering. *Tijdschrift voor Onderwijsresearch*, 11, 187-202.
- Vermunt, J.D.H.M., Lodewijks, J.G.L.C., & Simons, P.R.J. (1986). Hardop denken als onderzoeksmethode naar regulatieprocessen bij tekstbestudering. *Tijdschrift voor Onderwijsresearch*, 7, 187-202.
- Wouters, L., & Jong, T. de (1982). Hardop denken tijdens tekstbestudering. *Tijdschrift voor Onderwijsresearch*, 7, 60-75.

Manuscript aanvaard 18-9-1995

Auteurs

F.J. Prins en V.V. Busato waren ten tijde van dit onderzoek verbonden als toegevoegd docent en onderzoeker aan de Faculteit der Psychologie, Vakgroep Psychonomie, Universiteit van Amsterdam. Momenteel werken ze als gastonderzoeker bij dezelfde vakgroep.

C. Hamaker is als universitair hoofddocent verbonden aan de Faculteit der Psychologie, Vakgroep Psychonomie, Universiteit van Amsterdam.

K.H. Visser is werkzaam als hoofd van bureau Onderwijs en Onderzoek van de Faculteit der Psychologie van de Universiteit van Amsterdam.

Adres: Faculteit Psychologie, Vakgroep Psychonomie, Universiteit van Amsterdam, Roetersstraat 15, 1018 WB Amsterdam.

Abstract

A contribution to the validation of a part of the Inventory Learning Styles

F. J. Prins, V. V. Busato, C. Hamaker & K. H. Visser. *Pedagogische Studiën*, 1996, 73, 108-122.

One could ask if students study as they do according to the Inventory Learning Styles (ILS) of Vermunt (1992). With the ILS it is possible to create a learning style profile. The aim of this study is to make a contribution to the validation of a part of the ILS. 27 first year psychology students, 5 with a reproduction directed learning style, 8 with a meaning directed

learning style and 13 with an undirected learning style, were asked to think aloud while studying a text of ten pages about social psychology. The verbal data were divided into meaningful units and classified into (sub)categories. In agreement with the theory of Vermunt the reproduction directed students memorised more than the meaning directed and undirected students. For some subcategories the students with a meaning directed learning style tended to show more behaviour characteristic for their learning style. For other (sub)categories no differences were found between the learning styles, probably due to the small sample and the relative small text the students had to learn.

Bijlage A

Scoringscategorieën voor de hardop-denken-protocollen

1 Structureren

1.1 Niet-inhoudelijk: de lerende brengt structuren in de tekst aan, welke primair een niet-inhoudelijk karakter dragen.

1.1.1 Benoemen (SB): De lerende geeft aan een stuk tekst een benoeming die betrekking heeft op de (structurele) positie binnen een groter geheel.

Voorbeelden:

- "Dit is een vraagstelling"
- "Dat was een voorbeeld"
- "Paragraaf 3.1."
- "Tweede alinea."

1.1.2 Structureren (in engere zin)

a) **Impliciet: (SIM)** De lerende brengt alleen een uiterlijke structuur aan, die niet reeds in de tekst aanwezig is.

Voorbeeld:

- "Dat zijn vier punten"
- "Die komen uit paragraaf twee."

b) **Expliciet: (SEX)** De aangebrachte structuur wordt tevens expliciet ingevuld.

Voorbeeld:

- "Ik heb dus nu drie kolommen, één met ..., één met ..., etc"

1.1.3 Structurele Inferenties (SI): De lerende

doet voorspellingen of trekt conclusies over de opbouw van een tekstdeel.

Voorbeeld:

- "Waarschijnlijk volgt nu een beschrijving van de theorieën".

1.1.4 Oriënteren (SO). De lerende werkt een gedeelte van de tekst door op een globale manier, zodat een totaalbeeld van de tekst of een passage kan ontstaan.

Voorbeeld:

- "Dan krijgen we eerst Kelley en Thibaut, dan Mulder, Milgram en dan volgt de conclusie".
- "Hier wordt het experiment samengevat."

Relateren

1.2 Inhoudelijk: de lerende brengt structuren aan welke een primair inhoudelijk karakter dragen.

1.2.1 Kernwoorden (RKW). De lerende selecteert *letterlijk* uit een vrij groot stuk tekst (alinea, paragraaf) een woord (uitdrukking) dat kennelijk naar het oordeel van de lerende een centrale functie in die alinea of paragraaf vervult, in die zin dat het woord of de uitdrukking als 'kapstok' voor de overige betrokken informatie zou kunnen fungeren.

Voorbeeld: in het Milgram-experiment: "Gehoorzaamheid".

1.2.2 Kernzinnen (RKZ). Zoals bij RKW maar i.p.v. een woord wordt letterlijk een frase (zin, zinsdeel, twee opvolgende zinnen) geselecteerd. Alleen lidwoorden mogen eventueel weggelaten worden, anders is er sprake van herformuleren. De zin bevat een *werkwoord*, anders is er sprake van een *kernwoord*.

1.2.3 Vergelijken (RV). De lerende geeft aan dat er óf een verschil, óf een overeenkomst, óf een tegenspraak bestaat tussen twee of meer delen van de tekst. De lengte van deze delen kan variëren van zinnen tot paragrafen. De vergelijking staat niet expliciet in de tekst. (Belangrijk is het onderscheid met CVCH. Bij CVCH wordt er informatie aan de tekst toegevoegd).

Voorbeeld:

- "Mulder is het dus helemaal niet eens met Kelley en Thibaut"
- "Is hetzelfde als bij beloningsmacht..."
- "Is anders dan bij Kelly."

1.2.4 Integreren (RI). De lerende verbindt twee of meerdere delen tekst (alineas) zodanig dat er een eenheid in gescheiden aangeboden informatie ontstaat. Het proces RI omvat het proces RV, omdat voor er geïntegreerd kan worden er een vergelijking tussen de betrokken delen tekst gemaakt moet worden.

Voorbeeld:

- "Dus French en Raven onderscheiden o.a. deskundigheidsmacht, wat bij Milgram de macht van de proefleider over de proefpersoon is".

1.2.5 Overkoepelen (RO). De lerende abstrahert uit meerdere delen tekst (N.B. de delen kunnen ook zinnen zijn) dat element dat als kenmerkend voor die delen tekst kan gelden. Een dergelijk overkoepelend begrip staat niet als zodanig in de tekst. Het proces RO omvat het proces RI en RV. Deze drie processen vertonen een hiërarchische ordening.

Voorbeeld:

- "Het gaat hier dus eigenlijk steeds om afhankelijkheid".

2 Concreet verwerken

De lerende voegt informatie aan de tekst toe.

2.1 Voorbeeld geven (CVV). De lerende

geeft een voorbeeld, verklaring, of een definitie, niet gegeven in de tekst, voor een in de tekst genoemd begrip, principe, theorie etc.

Voorbeeld:

- "Zoals het bij Milgram ging, ging het met de kampbeulen in de tweede wereldoorlog dus ook".

2.2 Refereren (CVR). De lerende geeft aan dat de inhoud van de tekst verbanden heeft met al dan niet gespecificeerde externe informatie, die niet bedoeld is als voorbeeld, definitie of verklaring.

Voorbeeld:

- "In Atkinson heb ik ook zoiets gelezen".

2.3 Conclusies/Hypothesen stellen (CVCH).

Uit de tekst maakt de lerende afleidingen m.b.t. de inhoud in de vorm van hypothesen of conclusies.

Voorbeeld:

- Tekst: "Een tijd later werd de proefpersonen gevraagd of ze spijt hadden aan het experiment deelgenomen te hebben. Dit was bij slechts 1,3 % van de deelnemers het geval."

Protocol: "Dus het onderzoek was ethisch wel verantwoord".

- "Dus ongelijkheid komt voort uit onafhankelijkheid."

3 Kritisch verwerken

3.1 Evaluatie zonder uitleg (KVEZ). De lerende doet een uitspraak over het belang van de in de tekst gepresenteerde informatie zonder hieraan een inhoudelijke onderbouwing te geven.

Voorbeeld:

- "Dat is een open deur".
- "Hoofdpunt is..."
- "Goede definitie."

3.2 Evaluatie met uitleg (KVEM). Zoals KVEZ, waarbij de lerende hierbij een uitleg geeft d.m.v. een alternatief, een tegenvoorbeeld, etc. De uitleg kan inhoudelijk of niet-inhoudelijk zijn.

Voorbeeld:

- "Dat weet ik niet of dat wel zo is, het kan ook aan de experimentele condities liggen".
- "Dit is schuingedrukt en dus belangrijk."

4 Zelfsturing leerprocessen

4.1 Herformuleren

4.1.1 Woord herformuleren (ZLHW). De lerende geeft een synoniem, omschrijving of vertaling van een in de tekst genoemd woord of begrip, in eigen woorden en niet allemaal dezelfde woorden als in de oorspronkelijke tekst.

Voorbeeld:

- *"Referentiemacht is dus macht gebaseerd op identificatie".*

4.1.2 Zin herformuleren (ZLHZ). De lerende herformuleert een zin of een zinsdeel uit de tekst.

Voorbeeld:

Tekst: *"De machtsafstandreductietendens is sterker bij kleinere machtsafstand"*

Protocol: *"Je bent dus meer geneigd de machtsafstand te verkleinen t.o.v. diegenen die een paar stapjes boven je op de sociale ladder staan dan t.o.v. diegenen die buiten bereik zijn".*

4.1.3 Passage herformuleren (ZLHP). De lerende herformuleert een passage uit de tekst, zonder daarbij de inhoud of de structuur te veranderen. Er is sprake van een passage als er twee of meer hele zinnen herformuleerd worden.

4.1.4 Beeldmateriaal herformuleren (ZLHB). De lerende herformuleert in eigen woorden wat er in een tabel, figuur of grafiek staat.

Voorbeeld:

- *"Machtsbezit, blok eromheen, valt uiteen in bronnen en beschikking over middelen."*

4.2 Plannen (ZLP). Dit betreft het sturen van het leerproces op macro-niveau. De leerling geeft aan wat hij gaat doen om het gewenste eindresultaat te bereiken, zegt hoe hij de aandacht gaat verdelen over tekstonderdelen, geeft de inhoud en de volgorde aan van minstens twee successieve leeractiviteiten e.d. Het kan ook een opdracht aan zichzelf zijn. Je kan ook plannen om iets *niet* te doen. (Belangrijk om het te onderscheiden van ECEH. Bij ECEH gaat het om iets waar men mee bezig is of wat men al gedaan heeft).

Voorbeeld:

- *"Dat kijk ik even door..."*
- *"Ik lees die zin twee keer over."* (Je kunt niet bezig zijn iets twee keer over te lezen, dat moet wel toekomstig zijn).
- *"Opschrijven!"* (Opdracht aan jezelf om iets te gaan doen).
- *"Dit schrijf ik niet op."*

4.3 Zichzelf een vraag stellen (ZLV). Lerende stelt zichzelf een inhoudelijke vraag over de stof.

Voorbeeld:

- *"Ik zou wel willen weten wat het verschil is tussen referentiemacht en legitieme macht."*
- *"En hoe zit het dan met ...?"*

5 Memoriseren

De lerende herhaalt hardop bepaalde informatie uit de tekst of informatie eerder door hemzelf gegeven, of selecteert informatie die niet als kernwoord of kernzin gekenschetst kan worden.

5.1 Abstracte begrippen herlezen (MAB). Hierbij gaat het om informatie als begrippen en definities. Per begrip en per definitie scoren.

5.2 Concreet materiaal herlezen (MCM). Hierbij worden details of illustratief materiaal herhaald.

5.3 Ezelsbruggetjes (MEZ). Lerende bedenkt een niet-inhoudelijk trucje om rijtjes of feitjes uit het hoofd te leren, eventueel met uitleg.

Voorbeeld:

- *"B b d i l r, dat zijn de voorletters."*
- *"3 maal 2, dat is gemakkelijker te onthouden."*

6 Onbegrip constateren

6.1 Woord niet kennen (OWN). De lerende geeft aan een woord niet te kennen.

6.2 Zin of tekstdeel niet begrijpen (OZN). De lerende geeft aan een zin of tekstdeel niet te begrijpen.

6.3 Inhoudelijke uitleg vragen aan de proefleider (OU).

7 Exploratieve variabelen

7.1 Markeren & aantekenen

7.1.1 Markeren in de tekst (EMIT). De lerende onderstreept een woord, zin of passage

7.1.2 Aantekeningen naast de tekst (EANT). De lerende maakt aantekeningen in de kantlijn of op een kladblad.

7.2 Herlezen (EHRL). De lerende herleest een tekstdeel, zonder daarbij aan te geven wat zijn motivatie tot het herlezen is.

7.3 Constateren eigen handelingen (ECEH). De lerende beschrijft wat hij/zij doet of gedaan heeft.

7.4 Begrip constateren (EBC). De lerende constateert dat hij/zij een tekstdeel begrepen heeft.

7.5 Werkwijze (EW). Hieronder vallen opmerkingen van de proefpersoon die betrekking hebben op zijn studiemethoden.

7.6 Restcategorie (ERC). Overig.

Autoritarisme en etnocentrisme bij leerlingen uit lagere opleidingsniveaus: instemmingstendens of opinie?

W. A. M. Vollebergh

Samenvatting

In dit artikel wordt nagegaan of de hoge autoritarisme- en etnocentrisme-scores van leerlingen uit lagere opleidingsniveaus het gevolg zijn van een instemmingstendens. Hierbij werd gebruik gemaakt van ongebalanceerde en gebalanceerde autoritarisme- en etnocentrisme-items. Er werd geen steun gevonden voor de stelling dat deze hogere instemming het gevolg van een instemmingstendens is, noch voor de stelling dat de ambiguïteit van items tot een instemmingstendens zou leiden. De validiteit van de scores op de autoritarisme- en etnocentrisme-schaal werd nagegaan met behulp van de politieke partijvoorkeur. De relatie tussen etnocentrisme en politieke partijvoorkeur bleek eenduidig, binnen alle opleidingsniveaus, terwijl deze relatie bij autoritarisme moeilijker interpreteerbaar bleek. De conclusie luidde dat zich weliswaar een validiteitsprobleem bij de autoritarismeschaal voordoet, vooral bij leerlingen uit lagere opleidingsniveaus, maar dat dit niet als ondersteuning voor een instemmingstendens opgevat kan worden.

Het onderzoek werd gehouden in het kader van de landelijke scholierenverkiezingen. Er namen ruim 2000 leerlingen uit verschillende schooltypen aan deel.

Inleiding

Jeugdonderzoek in Nederland naar de waarden en normen van jongeren laat al meer dan een decennium zien, dat er belangrijke verschillen bestaan tussen de jongeren afkomstig uit de verschillende opleidingsniveaus. Als het gaat om opvattingen over gezag en over etnische minderheden, dan komen leerlingen uit lagere opleidingsniveaus (L/VBO, MAVO) in dit onderzoek regelmatig als veel autoritairder en veel etnocentrischer uit de bus (Vollebergh, Raaijmakers & Meeus, 1995). Er bestaan ver-

schillende verklaringen voor dit verschil. Een sociale verklaring legt de nadruk op de sociaal-culturele omstandigheden binnen en buiten de schoolsituatie. Jongeren uit lagere opleidingsniveaus zijn vaker afkomstig uit gezinnen in minder kansrijke sociaal-economische situaties, zij zijn om die reden banger voor statusverlies en voor concurrentie van allochtone Nederlanders en wijzen diens cultuur om die reden af (Eisinga & Scheepers, 1989). Een meer cognitieve verklaring legt de nadruk op verschillen in cognitieve en intellectuele ontwikkeling: jongeren uit hogere opleidingsniveaus beschikken over veel meer kennis over de politieke context, zij zijn daardoor in staat om op een abstracter niveau te redeneren en zien om die reden af van de simpele verklaringen voor maatschappelijk onrecht, zoals die in de autoritarisme- en de etnocentrisme-schaal geformuleerd worden (Raaijmakers, 1993). Een meer methodologische verklaring stelt dat jongeren uit lagere opleidingsniveaus meer dan jongeren uit hogere opleidingsniveaus de neiging zouden hebben om in te stemmen met de items, die onderzoekers hun voorleggen en om die reden hogere scores vertonen: hun hogere scores zouden een artefact zijn van de methode en niet wijzen op een verschil in opinie (Schuman, Bobo & Krysan, 1992).

Deze laatste verklaring is van een andere orde dan de eerste twee verklaringen, omdat zij het bestaan van 'werkelijke' opleidingsverschillen ontkent. Hoewel sinds het verschijnen van *De Autoritaire Persoonlijkheid* het bestaan van 'acquiescence' of 'response-set' (de neiging om steeds met de voorgelegde items in te stemmen, onafhankelijk van de inhoud daarvan) naar voren is gebracht, is hiervan in het Nederlands onderzoek bij adolescenten tot dusver niet veel te merken geweest. In dit onderzoek wordt vrijwel steeds gebruik gemaakt van ongebalanceerde autoritarisme- en etnocentrismeschalen. Dat maakt de resultaten van deze studies bijzonder kwetsbaar voor deze al-

ternatieve verklaring.

Discussies over het bestaan van een instemmingstendens werden vooral gevoerd in de jaren vijftig en zestig, na het verschijnen van de *Autoritaire Persoonlijkheid* (zie voor een overzicht Brown, 1965; Meloen, 1983). Een van de belangrijkste argumenten ten gunste van de instemmingstendens werd naar voren gebracht door Peabody (1961), die liet zien dat een relatief groot percentage van de instemmers met de autoritarisme-items ook instemt met hun negatie, terwijl dat bij andere schalen minder het geval was (bijvoorbeeld bij de door hem gebruikte conservatisme-schaal). Dit zou wijzen op een instemmingstendens als reactie op de ambiguïteit van de autoritarisme-items, die immers met opzet een beetje indirect en versluierend waren geformuleerd. Instemming hiermee zou daarom niet wijzen op een autoritaire stellingname, maar op het ontbreken van een opinie. Afwijzing van de items daarentegen zou wel op een valide wijze de afwijzing van een autoritaire stellingname meten. Anderen wezen op alle conceptuele en methodologische problemen bij het balanceren of omdraaien van items (zie vooral Christie, Havel & Seidenberg, 1958) en betwijfelden de veronderstelde ambiguïteit van de oorspronkelijke autoritarisme-items. Zo bleek de rangorde in de mate van instemming met oorspronkelijke autoritarisme-items opmerkelijk invariant over een tijd van twintig jaar (Samelson & Yates, 1967).

Er is geen sluitende conclusie uit deze discussie te trekken. Over het algemeen wordt aangenomen dat er wel zo iets als 'acquiescence' bestaat, en dat het daarom beter is om met gebalanceerde autoritarisme-schalen te werken, waarvan er verschillende in omloop zijn. Een van de meest recente versies daarvan is de RWA-schaal (Right-Wing Authoritarianism) van Altemeyer (1988), die ook in Nederlands onderzoek is uitgeprobeerd en gevalideerd (Meloen, 1991).

Een belangrijk nadeel van het empirisch onderzoek, dat in het kader van deze discussie werd uitgevoerd was, dat dit zich beperkte tot onderzoek bij studentenpopulaties. De generaliseerbaarheid van deze resultaten is daarmee beperkt, en een extrapolatie van deze discussie naar een interpretatie van opleidingsverschillen discutabel. Toch heropenden Schuman et al. in 1992 deze discussie met de stelling, dat

vooral bij respondenten uit lagere opleidingsniveaus de instemming met de F-schaal items op een instemmingstendens zou wijzen. Uit hun onderzoek bleek, dat bij deze respondenten een correlatie tussen autoritarisme-items en indices voor etnocentrisme ontbrak, terwijl deze bij respondenten uit hogere opleidingsniveaus wel aanwezig was. Zij concludeerden dat om die reden de hoge autoritarisme-scores van respondenten uit lagere opleidingsniveaus het resultaat van een instemmingstendens moesten zijn. Deze recente bijdrage zet weer vraagtekens bij het hoge autoritarisme (c.q. etnocentrisme) van lager opgeleide jongeren, vooral omdat de instemmingstendens ten gevolge van een gebrek aan opinie bij jongeren, met hun geringere kennis van de politiek dan volwassenen, nog sterker werkzaam zou kunnen zijn.

In dit artikel wordt verslag gedaan van een poging om de verklaring van de hogere instemming met autoritarisme en etnocentrisme bij jongeren uit lagere opleidingsniveaus in termen van een instemmingstendens nader te onderzoeken door het gebruik van gebalanceerde autoritarisme- en etnocentrisme-items. De vraag die we daarbij willen beantwoorden luidt: is de hogere instemming met autoritarisme en etnocentrisme bij jongeren uit lagere opleidingsniveaus het resultaat van een instemmingstendens, of wijst zij op een verschil in opinie? We beschikken op grond van het bovenstaande over drie mogelijke hypothesen omtrent de hogere instemming van jongeren uit lagere opleidingsniveaus:

1. De hogere instemming is het gevolg van een algemene instemmingstendens: jongeren uit lagere opleidingsniveaus hebben sterker dan jongeren uit hogere opleidingsniveaus de neiging om in te stemmen met items die hen in het kader van sociaal-wetenschappelijk onderzoek worden voorgelegd. Hun instemming geeft geen opinie weer.
2. De hogere instemming is een reactie op de ambiguïteit van items: als de inhoud van bepaalde items ambigu is, dan reageren vooral jongeren uit lagere opleidingsniveaus hierop door met het item in te stemmen. Hun instemming reflecteert daarom eveneens een gebrek aan opinie.
3. De hogere instemming is een reactie op de inhoud van de schaal: instemming komt

voort uit het aanhangen van een opinie, die met de items in de betreffende schaal goed wordt weergegeven.

1 Methode

1.1 Onderzoekspopulatie

Het onderzoek werd uitgevoerd in het kader van de scholierenverkiezingen in 1994. Aan deze verkiezingen namen in totaal 390.000 scholieren deel, afkomstig uit alle provincies in Nederland, maar met een oververtegenwoordiging van de provincies Gelderland (11,3%), Noord-Holland (14,8%), Zuid-Holland (19%) en Noord-Brabant (16%). Het totale aantal leerlingen op scholen die bereid waren aan verder onderzoek mee te werken was 125.000. Uit deze scholen werd at random een aantal scholen getrokken voor verder onderzoek. Via contactpersonen (doorgaans leraren maatschappijleer en/of geschiedenis) werden per school ook at random vier klassen geselecteerd. Van de 4000 verstuurd vragenlijsten werd ruim 50% geretourneerd ($n=2311$), afkomstig uit 30 verschillende scholen verspreid door het land. De deelnemende jongeren waren als volgt over de verschillende opleidingsniveaus verdeeld: 132 uit VBO-LBO, 530 uit MAVO, 285 uit MBO, 666 uit HAVO, en 698 uit het VWO. De hogere opleidingsniveaus zijn duidelijk beter vertegenwoordigd. Wel waren er ongeveer evenveel meisjes (48%) als jongens (52%) bij het onderzoek betrokken. De gemiddelde leeftijd van de leerlingen in het voortgezet onderwijs (excl. MBO) was 16 jaar. Omdat de gemiddelde leeftijd van de MBO-leerlingen ongeveer anderhalf jaar hoger was dan die van de overige leerlingen, werden hun resultaten in de analyses voor dit artikel niet meegenomen.

1.2 Instrumentarium

Autoritarisme werd gemeten met een veel gebruikte verkorte versie van de autoritarisme-schaal, bestaande uit 5 items (Raaijmakers, 1986). Deze schaal meet gehechtheid aan orde en gezag. Een voorbeeld: Het belangrijkste wat kinderen moeten leren is gehoorzaamheid en eerbied voor gezag. Deze schaal werd aangevuld met 5 gebalanceerde items. Vier hiervan waren ontleend aan de gebalanceerde autoritarisme-schaal van Altemeyer (1988),

maar deels geherformuleerd omdat de items van Altemeyer vaak samengestelde items zijn. Uit elk item werd een stelling geselecteerd om het item eenduidiger in betekenis te maken. Dat leek ons met name voor gebruik bij jongeren uit lagere opleidingsniveaus van belang. De gebalanceerde items waren respectievelijk: 1. Het is belangrijk om de rechten te respecteren van radicale actievoerders (Altemeyer: It is important to protect fully the rights of radicals and deviants). 2. Het is goed kritisch te zijn tegenover bestaande tradities en gewoonten in de samenleving (Altemeyer: students in high school and university must be encouraged to challenge their parents' ways, confront established authorities, and in general criticize the customs and traditions of our society). 3. Radicale, vooruitstrevende ideeën zijn nodig voor het leefbaar houden van de samenleving (Altemeyer: it is best to treat dissenters with leniency and an open mind, since new ideas are the lifeblood of progressive change). 4. Kinderen moeten zoveel mogelijk zelf uitmaken wat zij in hun leven belangrijk vinden (Altemeyer: it is wonderful that young people today have greater freedom to protest against things they don't like and to "do their own thing"). 5. Er is over het algemeen veel meer orde en regelmaat op scholen dan goed is voor jongeren (niet ontleend aan Altemeyer; dit item werd toegevoegd om het tegendeel te meten van de gehechtheid aan orde - een van de centrale kenmerken van autoritarisme).

Etnocentrisme werd gemeten met een etnocentrisme-schaal bestaande uit 7 items (Raaijmakers, 1986). Deze schaal meet een negatieve houding ten opzichte van etnische minderheden. Een voorbeeld: Buitenlanders vormen een bedreiging voor onze cultuur. Zij werd aangevuld met 5 items die juist een positieve houding ten opzichte van etnische minderheden meten. Drie items werden ontleend aan Kleinpenning (1993): 1. Etnische minderheden in Nederland zijn vrij om hun eigen manier van leven te handhaven. 2. Etnische minderheden moeten dezelfde politieke en sociale rechten hebben als de Nederlanders. 3. Het is goed dat er extra bijscholingscursussen zijn voor etnische minderheden. Twee items werden toegevoegd om hun actuele betekenis: 4. Nederland moet een gastvrij land blijven voor politieke vluchtelingen (asielzoekers). 5. Nederland

Tabel 1
Resultaten binnen de schalen

	Schaal											
	Autoritarisme						Etnocentrisme					
	Allen	LBO	MAVO	HAVO	VWO	Opleidingseffect	Allen	LBO	MAVO	HAVO	VWO	Opleidingseffect
I Scores op plus-versie												
Gemiddelde	3.34	3.60	3.50	3.35	3.16	ANOVA: F(3,1625)=31.12, p<.001	2.42	2.96	2.54	2.31	2.08	ANOVA: F(3,1482)=35.76, p<.001
SD	.62	.62	.56	.60	.64		.90	1.27	1.01	.82	.72	
Cronbach's alpha	.68	.63	.62	.67	.70		.91	.94	.91	.89	.88	
II Scores op min-versie												
Gemiddelde	3.17	3.28	3.23	3.10	3.18	ANOVA: F(3,1082)=3.81, p<.01	3.60	3.22	3.54	3.58	3.68	ANOVA: F(3,1640)=10.29, p<.001
SD	.54	.65	.57	.52	.51		.69	1.00	.72	.68	.61	
Cronbach's alpha	.41	.25	.45	.41	.43		.72	.79	.70	.72	.73	
III Scores op totaal-versie												
Gemiddelde	3.08	3.26	3.14	3.12	2.99	ANOVA: F(3,988)=10.27, p<.001	2.32	3.00	2.46	2.33	2.17	ANOVA: F(3,1303)=24.47, p<.001
SD	.43	.45	.40	.42	.44		.74	1.13	.82	.71	.61	
Cronbach's alpha	.55	.37	.47	.54	.60		.90	.93	.90	.90	.89	
V Correlatie (plus*min)	-.05 ns	-.01 ns	.05 ns	-.07 ns	-.13**		-.64***	-.55**	-.60**	-.66***	.65**	

moet een gastvrij land blijven voor mensen uit oorlogslanden, waar ook ter wereld.

Daarnaast werd aan alle leerlingen gevraagd om aan te geven, wat zij bij de laatste scholierenverkiezingen (die de week daarvoor waren gehouden) hadden gestemd.

1.3 Data-analyse

De schaal-items kenden 6 antwoordcategorieën: 1. helemaal mee eens; 2. mee eens; 3. niet eens/niet oneens; 4. niet mee eens; 5. helemaal niet mee eens. Om hoge scores te laten corresponderen met instemming met een bepaalde schaal, werden deze omgecodeerd. Als additionele antwoordcategorie werd toegevoegd: 6. nog nooit over nagedacht. Deze laatste categorie werd toegevoegd om het lukraak kiezen van een antwoordcategorie te vermijden en enkel echte 'opinions' te meten. Het toevoegen van deze antwoordcategorie heeft, mits de 6 als 'missing' wordt gecodeerd, als nadeel een sterke verhoging van het aantal missings per item, en per saldo per schaal. Voor de vraagstelling van dit artikel biedt deze antwoordcategorie echter een goede operationalisatie van het ontbreken van een opinie over een bepaald item.

Bij de analyse van de antwoorden gingen we eerst alle gebruikelijke indicaties na, dat wil zeggen: gemiddelde scores en standaarddeviaties, betrouwbaarheidscoëfficiënten, en correlaties tussen plus- en minversies van de schalen (zie Tabel 1). Alvorens de betrouwbaarheidscoëfficiënten te berekenen werd factoranalyse uitgevoerd (principale componenten-analyse met varimax rotatie, listwise deletion) over alle items van een schaal. Bij de etnocentrisme-items werden weliswaar twee factoren met een eigenwaarde >1 geëxtraheerd, maar de etnocentrisme-items laadden allen >.50 op de eerste (ongeroteerde) factor. Bij autoritarisme werden meerdere factoren met een eigenwaarde >1 geëxtraheerd (bij LBO 5 factoren, bij MAVO en HAVO 3 factoren en bij VWO 4 factoren). Rotatie met twee factoren leverde bij de LBO-leerlingen geen goed interpreteerbare factor-matrix op, bij de overige leerlingen wel: in alle gevallen hadden de items van de plus-versie een lading van >.50 op de eerste geroteerde factor, terwijl de items van de min-versie een lading van >.35 op de tweede geroteerde factor hadden. Eliminatie van items leverde geen be-

	Antwoorden op plus-versie	
Antwoorden op min-versie	<i>Instemming</i>	<i>Afwijzing</i>
<i>Instemming</i>	Dubbele instemming	Anti-inhoud
<i>Afwijzing</i>	Pro-inhoud	Dubbele afwijzing

NOOT: instemming/afwijzing betekent dat bij 3 uit 5 items van de schaal (4 uit 7) het antwoord instemmend (4,5) c.q. afwijzend (4,5) is.

Figuur 1. Antwoordpatronen op schalen in plus- en minversies (cf. Peabody, 1961)

tere resultaten, en geen hogere betrouwbaarheidscoëfficiënt van de plus- of min-schaal.

Vervolgens berekenden we het percentage respondenten met een bepaald antwoordprofiel. Daarbij werd als volgt te werk gegaan (Peabody, 1961).

We gingen er van uit dat er van een instemmende opinie ten aanzien van een (sub)schaal sprake is als minimaal 3 van de 5 (4 van de 7, bij etnocentrisme) items instemmend beantwoord worden (dat wil zeggen antwoordcategorie 4 of 5), en van een afwijzende opinie als minimaal 3 van de 5 (4 van de 7) items afwijzend beantwoord zijn (antwoordcategorie 1 of 2). Een consistente keuze vóór de inhoud van een bepaalde schaal bestaat dan uit de combinatie van instemming met de plus-versie en afwijzing van de min-versie; een consistente keuze anti-inhoud daarentegen uit afwijzing van de plus- en instemming met de min-versie. Daarnaast bestaan twee antwoordpatronen, die tot dusver in onderzoek in termen van een instemmings-, c.q. een afwijzingstendens geïnterpreteerd werden: dubbele instemming, of dubbele afwijzing. Zowel het antwoord 'niet eens/niet oneens' (3), als het antwoord 'nog nooit over nagedacht' (6) werd hierbij als het ontbreken van een opinie geïnterpreteerd. Dat leek ons een valide aanname. Zouden we dat niet doen, dan zouden we immers een autoritaire, c.q. niet autoritaire stellingname toeschrijven aan respondenten op grond van slechts 1 of 2 instemmende items (uit 5). Bovendien is het juist voor de hier te presenteren analyse belangrijk om onderscheid te maken tussen respondenten met een duidelijke mening, en respondenten waar deze ontbreekt. Het aantal ontbrekende antwoordpatronen is door deze procedure bijzonder groot, vooral bij de autoritarisme-schalen (de min-versie). Ondanks het ontbreken van

een sterke gebalanceerde autoritarismeschaal presenteren we voor de volledigheid de gegevens met betrekking tot beide schalen. Bij hiernavolgende analyses op itemniveau komen we hier op terug.

Per item werd vervolgens nagegaan welk percentage repondenten een 3, respectievelijk een 6 scoorde. Vervolgens werd voor alle items binnen een subschaal het gemiddeld percentage van deze antwoordcategorie berekend. Ook dit werd voor de verschillende opleidingsniveaus nagegaan. Deze cijfers zijn te vinden in deel II en III van Tabel 2.

2 Resultaten

2.1 Het balanceren van de autoritarisme- en etnocentrisme-schaal

De betrouwbaarheidscoëfficiënten van de verschillende subschaaltjes laten zien, dat de positieve items een hogere betrouwbaarheid te zien geven dan de negatieve items. Dit geldt ook voor de etnocentrisme-items, maar hier is de alpha van de negatieve items ruimschoots binnen de grenzen van het aanvaardbare. Dat is bij autoritarisme niet het geval. De betrouwbaarheid van de positieve items is al aan de lage kant, maar de betrouwbaarheid van de schaal als geheel en van de negatieve items is laag. Enkel bij de leerlingen uit het VWO vinden we redelijke betrouwbaarheid van de gebalanceerde autoritarismeschaal, en een negatieve correlatie tussen de plus- en de min-versie, maar ook bij hen is de correlatie erg laag.

Deze gegevens zijn in overeenstemming met de vaak gerapporteerde problemen bij het construeren van een gebalanceerde autoritarisme-schaal. Ze vertonen veel overeenkomst met de bevindingen van Peabody (1961)¹. Ze bevestigen voorts Schuman et al.'s (1992) resultaten, die immers ook constateer-

Tabel 2
Antwoordpatronen bij verschillende opleidingsniveaus

	Schaal										
	Autoritarisme						Ethocentrisme				
	Allen	LBO	MAVO	HAVO	VWO	Opleidingseffect	Allen	LBO	MAVO	HAVO	VWO
I 'Antwoordcombinaties' (in %)											
'Pro-inhoud'	4.6	5.3	5.7	4.8	3.6	$\chi^2(12)=66.87$,	4.0	11.4	5.7	3.5	2.0
'Dubbele instemming'	11.6	18.9	14.9	10.4	9.0	$p<.001$	2.2	5.3	3.2	2.4	0.6
'Anti-inhoud'	4.2	2.3	2.1	2.6	7.9		44.3	16.7	33.8	47.4	54.4
'Dubbele afwijzing'	1.5	0	0.2	1.8	2.6		1.5	2.3	0.8	1.7	1.9
Anders	77.9	73.5	77.2	80.5	76.9		48.0	64.4	56.6	45.0	41.1
II 'Niet-opinie-antwoorden' plusversie											
(gemiddeld % per item)						Kendall:					Kendall:
Niet eens/niet oneens(3)	29.0	19.4	28.4	30.2	30.8	$\chi^2(3)=10.69$, $p<.01$	21.4	18.7	23.1	23.1	18.9
Nog nooit over nagedacht (6)	5.6	11.8	4.6	4.6	5.4	$\chi^2(3)=12.21$, $p<.01$	5.4	12.0	5.6	4.7	4.8
Totaal	34.8	31.2	33.0	34.8	36.2	$\chi^2(3)=5.94$, ns	26.8	30.7	28.7	27.8	23.7
III 'Niet-opinie-antwoorden' minversie											
(gemiddeld % per item)						Kendall:					Kendall:
Niet eens/niet oneens(3)	28.2	20.2	27.4	29.8	29.0	$\chi^2(3)=7.69$, $p<.05$	22.2	18.8	23.4	23.1	18.9
Nog nooit over nagedacht(6)	15.8	30.0	16.6	15.0	13.2	$\chi^2(3)=11.94$, $p<.01$	5.6	22.2	8.4	3.0	2.6
Totaal	44.0	50.2	44.0	44.8	42.2	$\chi^2(3)=8.26$, $p<.05$	27.8	41.0	31.8	24.4	25.6

den dat de samenhang in de verwachte richting bij respondenten uit de hogere opleidingsniveaus sterker was. Dit wordt bevestigd door het feit dat enkel bij de jongeren uit de hogere opleidingsniveaus (HAVO, VWO) significante correlaties worden aangetroffen tussen autoritarisme en etnocentrisme (zowel tussen de plus-versies - pearson correlatie respectievelijk .29 en .25, $p < .01$ - als tussen de totaal-versies - pearson correlaties respectievelijk .20 en .24, $p < .01$; geen bijzonder hoge correlaties overigens), terwijl deze bij de jongeren uit de lagere opleidingsniveaus ontbreken. De vraag is nu vooral, of we ook aanwijzingen kunnen vinden, dat de hogere instemmingspercentages bij jongeren uit lagere opleidingsniveaus om die reden ook als het effect van een instemmingstendens verklaard mogen worden. We bespreken onze drie hypothesen.

2.2 Hogere instemming bij jongeren uit lagere opleidingsniveaus door instemmingstendens

Een algemene instemmingstendens bij lager opgeleide jongeren zou moeten blijken uit een relatief hoog percentage dubbele instemmers bij zowel de autoritarisme als de etnocentrisme-schaal. Immers, als respondenten de neiging hebben om de onderzoeker te volgen, dan zou dit zich bij beide schalen moeten manifesteren. Een dergelijke instemmingstendens is echter niet aantoonbaar. Bij etnocentrisme is de instemming van leerlingen uit lagere opleidingsniveaus groter bij de plus-versie, maar hun afwijzing is veel groter bij de min-versie. Wel geldt dat het opleidingseffect bij de plus-versie groter is (het verschil tussen de LBO-leerlingen en de VWO-leerlingen uitgedrukt in gemiddelde *SD* is bij de plusversie ongeveer anderhalf keer zo groot). Dat zou erop kunnen wijzen dat instemmen 'makkelijker' is dan afwijzen bij jongeren uit lagere opleidingsniveaus - maar het zou er overigens ook op kunnen wijzen dat afwijzen voor jongeren uit hogere opleidingsniveaus makkelijker is dan instemmen. Het percentage dubbele instemmingen (of dubbele afwijzingen) is bij de etnocentrismeschaal echter laag. Overigens zijn de percentages dubbele instemmingen/afwijzingen bij de lager opgeleide jongeren wel systematisch ietwat hoger, maar bij etnocentrisme domineren ook bij hen de 'opinie-antwoorden' (28.1% versus 7.6%).

Bij etnocentrisme is dus geen substantiële instemmingstendens aanwijsbaar. Het feit dat we bij autoritarisme zoveel dubbele instemming aantreffen is daarom niet terug te voeren op een generale instemmingstendens. We kunnen de eerste hypothese op grond van onze bevindingen met de etnocentrismeschaal verwerpen.

2.3 Hogere instemming bij jongeren uit lagere opleidingsniveaus als reactie op ambiguïteit van items

De autoritarisme-items zouden het door hun ambiguïteit moeilijk kunnen maken om hierover een mening te vormen. Ondanks een gebrek aan opinie zouden de respondenten dan toch met deze items instemmen. Dat zou de grotere instemming bij leerlingen uit lagere onderwijsniveaus kunnen verklaren. Het ontbreken van een opinie is in ons onderzoek na te gaan met behulp van antwoordcategorie 6 (nog nooit over nagedacht) en antwoordcategorie 3 (niet eens/niet oneens).

Aan de percentages 'niet-opinie-antwoorden' in sector II en III van Tabel 2 is te zien dat de etnocentrisme-items het duidelijkst een opinie oproepen en wel voornamelijk een afwijzende. De autoritarisme-items roepen veel meer 'niet-opinie-antwoorden' op, vooral de min-versie. Binnen elk opleidingsniveau wordt bij de min-items van autoritarisme het grootste percentage niet-opinie-antwoorden gegeven (significant bij $p < .01$). Wel doet zich hierin, als te verwachten was, een opleidingseffect voor. Hoe lager de opleiding, hoe groter het aantal 'niet-opinie-antwoorden' (behalve bij de plus-versie van de autoritarisme-schaal). Dat is vaker in onderzoek aangetoond (Billiet, in druk). Daarnaast is er ook een duidelijk opleidingseffect in de keuze voor een bepaald 'niet-opinie-antwoord'. Jongeren uit lagere opleidingsniveaus zeggen vaker dat zij ergens nog nooit over hebben nagedacht, terwijl jongeren uit hogere opleidingsniveaus er vaker voor kiezen om hun mening in het midden te laten (niet eens/niet oneens).

Maar is er een relatie tussen een gebrek aan opinie en *instemming*, zoals zo vaak gesuggereerd is? Op twee manieren is dit geanalyseerd: a. door na te gaan of er een relatie is tussen de 'moeilijkheidsgraad' van een item (geoperationaliseerd als het percentage 'niet-opinie-

antwoorden') en de score op dit item bij de respondenten die op dit item wel een opinie (dat wil zeggen: 1, 2, 4 of 5) hebben weergegeven. Zou gebrek aan opinie tot instemming leiden, dan zou de instemming het hoogst moeten zijn bij items die het hoogste percentage 'niet-opinie-antwoorden' oproepen. Dit blijkt niet het geval. Bij autoritarisme ligt de relatie precies omgekeerd, en wel enkel bij de respondenten uit de lagere opleidingsniveaus. Bij jongeren uit het VBO/LBO is de correlatie tussen de percentages 'niet-opinie-antwoorden' per item en de itemscores van de jongeren die wel een opinie gaven $-.80$, significant bij $p < .01$ (indien als 'niet-opinie-antwoorden' enkel de 'nog nooit over nagedacht' categorie gerekend wordt $-.69$, significant bij $p < .05$). Dat wil zeggen: naarmate het percentage 'niet-opinie-antwoorden' groter is, is de gemiddelde instemming met het item bij de overige leerlingen lager. Bij de MAVO-leerlingen zijn deze correlaties respectievelijk $-.80$ ($p < .01$) en $-.68$ ($p < .05$), bij de HAVO-leerlingen $-.67$ ($p < .05$) en $-.51$ (n.s.), en bij de VWO-leerlingen $-.26$ (n.s.) en $-.20$ (n.s.). Voor zover we hier dus een relatie vinden, vinden we die bij de jongeren uit de lagere opleidingsniveaus en deze is tegengesteld aan de verwachtingen: items, waarover leerlingen zich moeilijker een mening vormen, worden eerder afgewezen. Bij de etnocentrisme-schaal zijn deze correlaties binnen elk opleidingsniveau afwezig.

Ook gingen we een mogelijke relatie tussen het ontbreken van een opinie en een instemmingsstendens na door te bezien of respondenten die op een bepaalde schaal een hoog percentage 'niet-opinie-antwoorden' hebben gegeven, op de overige items van die betreffende schaal sterker instemmen dan respondenten met een laag percentage niet-opinie-antwoorden. We gingen dit zowel na voor het geven van antwoord 6 (nog nooit over nagedacht) als voor het geven van een 'niet-opinie antwoord' in meer algemene zin (3 of 6). Hier toe staat ons uiteraard enkel analyse op itemniveau ter beschikking. We correleerden de score op een bepaald item met het aantal niet-opinie-antwoorden op de overige items van de betreffende schaal. We beperken ons in deze bespreking tot de correlaties tussen de 'niet opinie-score' van respondenten op een totale schaal (plus + min-versie) en hun score op items van

die schaal, waarbij zij wel een opinie hebben aangegeven. Dat betreft dus 10 (autoritarisme) en 12 (etnocentrisme) is 22 correlaties per opleidingsniveau. De veronderstelling, dat een gebrek aan opinie leidt tot sterkere instemming bij de items van een schaal wordt enkel bevestigd voor respondenten uit de hogere opleidingsniveaus. Bij LBO-leerlingen vinden we slechts 2 significante correlaties, een bij autoritarisme ($-.29$, $p < .01$) en een bij etnocentrisme ($-.18$, $p < .01$), die overigens beide negatief zijn; dat wil zeggen: hoe vaker geen opinie op de overige items van een schaal, hoe minder instemming met het betreffende item. Bij MAVO-leerlingen zijn geen correlaties bij autoritarisme, bij etnocentrisme zeven significant positief - (tussen $.10$ en $.13$, $p < .05$; een relatie tussen niet-opinie-score en instemming). Bij HAVO-leerlingen zien we een significantie correlatie bij autoritarisme ($.10$, $p < .01$) en zeven significante correlaties bij etnocentrisme (tussen $.12$ en $.19$, allen significant bij $p < .01$) en ook positief, dus wijzend op grotere instemming bij meer niet-opinie-antwoorden). Bij de VWO-leerlingen is er een significante negatieve correlatie bij autoritarisme ($-.09$, $p < .05$) en zes etnocentrisme-items een positieve correlatie (tussen $.10$ en $.17$, significant bij $p < .01$). Dit patroon wordt nog sterker zichtbaar als we zowel antwoordcategorie 3 (niet eens/niet oneens) als 6 (nog nooit over nagedacht) tot de 'niet-opinie-antwoorden' rekenen. Bij LBO-leerlingen vinden we dan nog steeds slechts 2 significante correlaties, die overigens beide negatief zijn (dat wil zeggen: hoe vaker geen opinie op de overige items van een schaal, hoe minder instemming met het betreffende item). Bij MAVO-leerlingen zijn twee correlaties bij autoritarisme significant negatief, maar bij etnocentrisme 7 significant positief (een relatie tussen niet-opinie-score en instemming). Bij HAVO-leerlingen zijn alle correlaties bij etnocentrisme nu significant (tussen $.14$ en $.44$, negen van de twaalf correlaties zijn $> .25$, allen significant bij $p < .01$) en positief, dus wijzend op grotere instemming bij meer niet-opinie-antwoorden. Bij de VWO-leerlingen is het patroon het sterkst aanwezig: drie autoritarisme items laten een positieve correlatie zien, en alle etnocentrisme-items (allen significant bij $p < .01$, slechts twee correlaties zijn $< .25$, zes correlaties zijn zelfs $> .30$). Voor jongeren uit

hogere opleidingsniveaus geldt dus dat hun instemming met etnocentrisme-items sterker is naarmate zij zich op de overige items van de etnocentrisme-schaal vaker van een opinie onthouden.

Deze resultaten wijzen er niet op dat de instemming van leerlingen uit lagere onderwijssoorten verklaard wordt door hun neiging sterker in te stemmen met items naarmate zij het moeilijker vinden om zich een opinie te vormen. Voor zover we een dergelijke neiging aantreffen is dat juist bij de leerlingen uit de hogere onderwijstypen het geval. De hogere instemming van jongeren uit lagere opleidingsniveaus laat zich dus door dit effect niet verklaren. We kunnen ook de tweede hypothese afwijzen.

2.4 Hogere instemming bij jongeren uit lagere opleidingsniveaus door een verschil in opinie

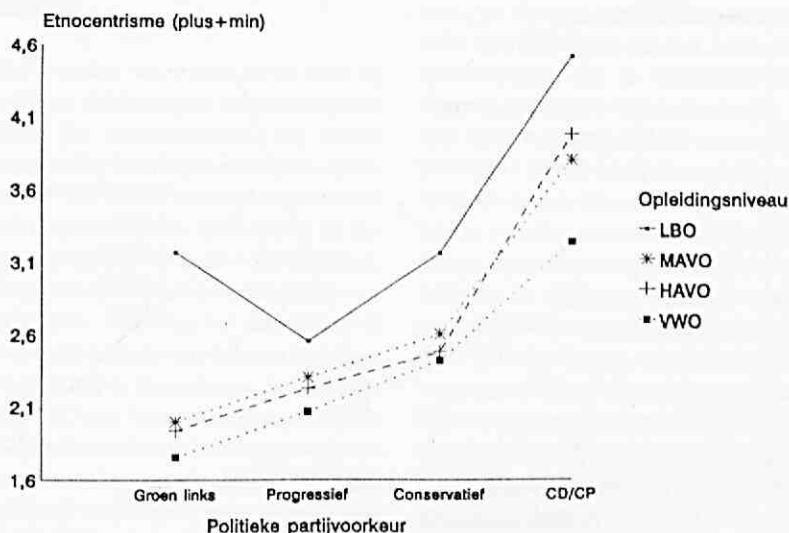
Als we aannemelijk willen maken dat de relatief hogere scores van leerlingen uit lagere opleidingsniveaus op de etnocentrisme- en de autoritarisme-schaal ontstaan omdat zij een andere opinie hebben dan jongeren uit hogere opleidingsniveaus, dan moeten we kunnen laten zien dat deze items op een betekenisvolle manier samenhangen met andere indicaties voor deze opinie.

Een van de belangrijkste indicatoren, waarover wij in dit scholierenonderzoek beschik-

ken, is de politieke partijvoorkeur van jongeren. Het onderzoek vond plaats in het kader van de scholierenverkiezingen. Wellicht om die reden vinden wij in dit onderzoek ook een bijzonder hoog percentage jongeren met een politieke partijvoorkeur (vaak is deze bij ongeveer een derde van de respondenten niet aanwezig; in dit onderzoek vinden we een percentage van 91.4% leerlingen met een partijvoorkeur). Een bijkomend voordeel van een analyse van partijvoorkeur is, dat deze niet lijkt op de Likert-items waaruit de verschillende schalen bestaan – er is geen sprake van meer of minder, er is geen sprake van instemming of afwijzing, er is enkel sprake van een keuze uit bepaalde alternatieven. Zoiets als instemmingstendens is daarmee bij het aangeven van een politieke partijvoorkeur niet aan de orde.

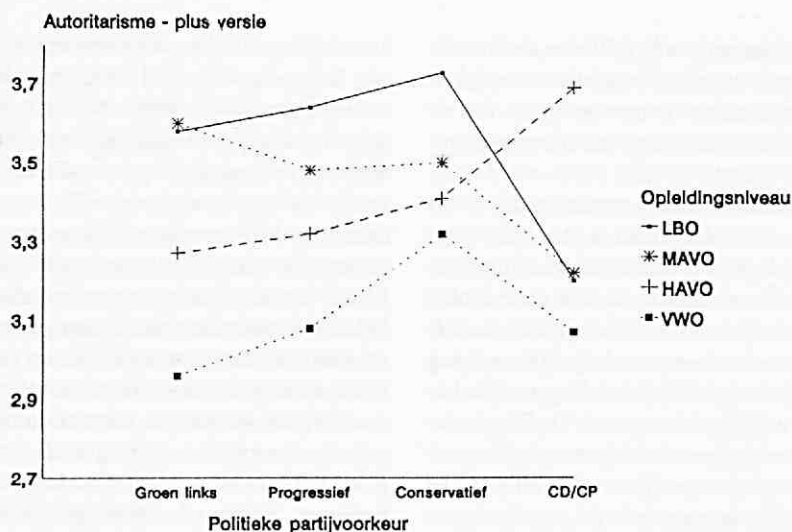
De relatie tussen de etnocentrisme-schalen en politieke partijvoorkeur is onmiskenbaar. Instemming met etnocentrisme neemt sterk toe naarmate de partijvoorkeur rechtser is. Dat geldt zowel voor de plus-versie als voor de min-versie van etnocentrisme, en het geldt voor respondenten binnen alle opleidingsniveaus. Dat bevestigt de validiteit van de etnocentrisme-scores. We zien dit geïllustreerd in Figuur 2.

Het verband tussen partijvoorkeur en autoritarisme is echter veel minder eenduidig. Bij de



NOOT: Resultaten ANOVA: hoofdeffect Partijvoorkeur $F(3,1192)=87.33, p<.001$, Opleidingsniveau $F(3,1192)=16.66, p<.001$, 2-Weg interactie $F(9,1192)=1.19, ns$.

Figuur 2. Etnocentrisme (totaal-versie) per partijvoorkeur en opleidingsniveau



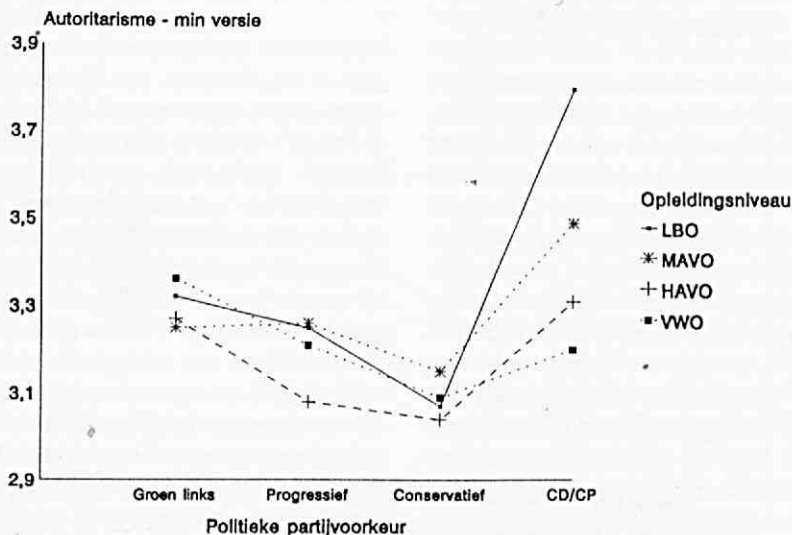
NOOT: Resultaten ANOVA: hoofdeffect Partijvoorkeur $F(3,1480)=6.04$, $p<.001$, Opleidingsniveau $F(3,1480)=32.73$, $p<.001$, 2-Weg interactie $F(9,1480)=2.39$, $p<.01$.

Figuur 3. Autoritarisme (plus-versie) per partijvoorkeur en opleidingsniveau

plus-versie van autoritarisme doet zich een interactie-effect tussen opleiding en politieke partijvoorkeur voor. Dit interactie-effect is al eerder in onderzoek aangetroffen (Vollebergh, 1991). Het wordt veroorzaakt door het feit dat bij de leerlingen uit de lagere opleidingsniveaus geen verband tussen autoritarisme en partijvoorkeur bestaat, en bij de leerlingen uit de hogere niveaus wel. Dit wordt geïllustreerd in Figuur 3. Afzonderlijke analyses laten zien,

dat in dit onderzoek zelfs enkel bij de VWO-scholieren het verband tussen autoritarisme en partijvoorkeur significant is.

De relatie tussen de min-versie van autoritarisme en politieke partijvoorkeur maakt een nog onverwachter fenomeen zichtbaar: extreem rechts (CD/CP) stemt het sterkst met de min-versie van autoritarisme in. Dit effect van politieke partijvoorkeur is zelfs sterker dan het



NOOT: Resultaten ANOVA: hoofdeffect Partijvoorkeur $F(3,995)=8.06$, $p<.001$, Opleidingsniveau $F(3,995)=3.55$, $p<.001$, 2-Weg interactie $F(9,995)=0.56$, ns.

Figuur 4. Autoritarisme (min-versie) per partijvoorkeur en opleidingsniveau

effect van opleidingsniveau. Analyse op item-niveau maakt zichtbaar dat drie items verantwoordelijk voor dit effect zijn (getoetst met ANOVA, hoofdeffecten significant bij $p < .05$): 1. Het is belangrijk om de rechten te respecteren van radicale actievoerders. 2. Radicale, vooruitstrevende ideeën zijn nodig voor het leefbaar houden van de samenleving. 3. Er is over het algemeen veel meer orde en regelmaat op scholen dan goed is voor jongeren. Bij één item doet zich alleen een opleidingseffect voor: hoger opgeleide jongeren stemmen hier sterker mee in (Het is goed kritisch te zijn tegenover bestaande tradities en gewoonten in de samenleving), bij één item is enkel een effect van politieke partijvoorkeur: de hoogste score bij extreem rechts (Kinderen moeten zoveel mogelijk zelf uit maken wat zij in hun leven belangrijk vinden).

Deze resultaten maken zichtbaar dat instemming met deze items niet het gevolg is van een instemmingstendens. Het lijkt er meer op dat we bij deze jongeren een geslaagde linguïstische herinterpretatie van deze items op het spoor zijn gekomen, die aan de als anti-autoritair bedoelde Altemeyer-items een extreem-rechtse verzetsdimensie geeft. Figuur 4 maakt dit aanschouwelijk.

3 Conclusie

In dit artikel vonden wij weinig steun voor de veronderstelling dat de hogere instemming met autoritarisme en ethnocentrisme, die steeds wordt gevonden bij leerlingen uit lagere opleidingsniveaus, op een instemmingstendens wijst en niet op een opinie. Ook wordt de instemming van jongeren uit lagere opleidingsniveaus niet opgeroepen door de ambiguïteit van een bepaald item: naarmate het moeilijker is om over een bepaald item van de autoritarisme-schaal een mening te formuleren, is hun neiging om met dit item in te stemmen juist kleiner. Wel blijkt dat naarmate leerlingen zich vaker van een mening op de items van een bepaalde schaal onthouden, hun neiging om met de overige items in te stemmen sterker is, maar dit effect vonden we enkel bij de ethnocentrisme-schaal en enkel bij leerlingen uit de hogere opleidingsniveaus. De hogere instem-

ming van de leerlingen uit de lagere opleidingsniveaus laat zich hierdoor niet verklaren.

Onze uitkomsten wezen er verder op dat het in het onderzoek met dit soort schalen – en wellicht in het bijzonder als in het onderzoek gebruik wordt gemaakt van jonge respondenten – bijzonder belangrijk is om nauwkeurig na te gaan welke betekenis de respondenten aan de verschillende items geven. In dit onderzoek bleken als anti-autoritair bedoelde items bijvoorbeeld een affiniteit met rechts-extreme radicaliteit te hebben. Ook de historische context mag daarbij niet uit het oog worden verloren. Zo is het de vraag of het begrip 'radicaal' in de afgelopen 10-15 jaar niet meer in het algemeen van betekenis is veranderd, nu extreem links nauwelijks nog de aandacht trekt, en extreem rechtse actiegroepen juist wel.

De volgende kanttekeningen moeten bij onze analyse geplaatst worden. In de eerste plaats bleek ook in ons onderzoek de autoritarisme-schaal bijzonder moeilijk te balanceren. Ondanks het feit dat wij enkele aangepaste versies van Altemeyer-items hebben gebruikt, zijn wij er niet in geslaagd een adequate gebalanceerde autoritarisme-schaal te ontwerpen. Enkel bij de VWO-leerlingen kwamen de resultaten enigszins overeen met de bedoelingen (dat wil zeggen: een redelijke betrouwbaarheid van de gebalanceerde schaal, en een significant negatieve correlatie tussen de plus- en de min-versie). Dat werpt nogmaals licht op alle risico's aan het werken met meetinstrumenten, die in onderzoek bij (Amerikaanse) studenten ontworpen zijn. Deze zijn niet zonder meer geschikt voor onderzoek bij leerlingen in het middelbaar onderwijs in Nederland. In het bijzonder lijken zij weinig geschikt voor het onderzoek bij leerlingen uit de lagere opleidingsniveaus. Zo kon gemiddeld 50% van de VBO-leerlingen geen opinie geven over de gebalanceerde autoritarisme-items (bij de VWO-leerlingen was dit overigens ook meer dan 40%). Daarmee is overigens niet gezegd dat dit verklaart waarom hun instemming met deze items groter was, of dat deze instemming als betekenisloos moet worden beschouwd. Onze resultaten wezen niet in deze richting. Wel moesten wij constateren dat de opinie, die met deze items gemeten wordt, niet overeenstemt met hun meetpretentie: vooral de leerlingen met extreem rechtse politieke voor-

keuren bleken met deze items in te stemmen. Ook bij de positieve autoritarisme-items ontbrak een relatie met politieke partijvoorkeur, behalve bij de VWO-leerlingen. Dit is overigens al vaker aangetoond (Vollebergh, 1991). Noch de (verkorte) autoritarisme-schaal, noch de door ons gebruikte items van de RWA (Right-Wing-Authoritarianism)-schaal konden dus in ons onderzoek hun meetpretentie waarmaken. Dat wijst op een serieus validiteitsprobleem van de autoritarisme-schaal. In dit artikel hebben wij echter willen aantonen dat dit validiteitsprobleem geen ondersteuning voor de aanname van een instemmingstendens bij lager opgeleide respondenten vormt, zoals Schuman et al. (1992) stelden.

Een tweede kanttekening betreft de wijze waarop de ambiguïteit van items in dit onderzoek werd gemeten. Hiertoe was de antwoordcategorie 'nog nooit over nagedacht' aan elk item toegevoegd. Het ontbreken van een opinie kon dus expliciet worden weergegeven. Onze vooronderstelling was dat de inhoud van een bepaald item meer ambigu, of meerduidelijk, was naarmate meer leerlingen moeite hadden met het bepalen van een standpunt. Dat leek ons een valide aanname. Het probleem is echter, dat doorgaans in onderzoek deze antwoordcategorie ontbreekt, en daarmee rijst de vraag in hoeverre we onze resultaten kunnen extrapoleren naar de resultaten van ander onderzoek. De centrale vraag is daarbij dan niet alleen of de respondenten, die in ons onderzoek bekenen dat zij nog nooit ergens over hebben nagedacht, in ander onderzoek wél een opinie zouden hebben weergegeven, maar ook of deze opinie dan meer instemmend van karakter zou zijn. De eerste vraag is moeilijk te beantwoorden. Op de tweede vraag kunnen wij een tentatief antwoord formuleren. Over het algemeen neemt het percentage 'nog nooit over nagedacht' antwoorden toe naarmate het opleidingsniveau lager is. Juist bij jongeren uit lagere opleidingsniveaus konden wij echter niet vaststellen dat het ontbreken van een opinie bij items van een schaal gepaard zou gaan met instemming met items, die wel beantwoord waren. Bij de leerlingen uit de hogere opleidingsniveaus was er wel een tendens tot sterkere instemming met items naarmate zij minder over andere items van de schaal hadden nagedacht. Deze tendens deed zich alleen bij de etnocentrisme-schaal

voor. Bij de autoritarismeschaal was er ook bij hen geen effect. Voor zover het meer gedwongen weergeven van opinies dus tot grotere instemming leidt, is dat bij jongeren uit de hogere opleidingsniveaus het geval, en niet bij jongeren uit lagere opleidingsniveaus, bij wie doorgaans juist wel de aanwezigheid van een instemmingstendens verondersteld wordt.

Noot

1. Peabody (1961) vindt bij de min-versie van zijn autoritarisme-schaal ook alpha's van rond de .40. Zijn resultaten suggereren dat eenzelfde factorstructuur (twee orthogonale factoren) aan de instemming met de autoritarisme-items ten grondslag zou kunnen liggen, maar hij rapporteert geen resultaten van factoranalyse.

Literatuur

- Altemeyer, B. (1988). *Enemies of freedom*. San Francisco: Jossey-Bass Publishers.
- Billiet, J., Carton, A. (in press). Do unbalanced scales affect the respondent's opinion? A plea for critical tests. In *Proceedings of the Conference on Survey Measurement and Data Quality*, Bristol, U.K., april 1995.
- Brown, R. (1965). The authoritarian personality and the organization of attitudes. In R. Brown, *Social Psychology* (pp. 477-546). New York: Free Press.
- Christie, R., Havel, J., & Seidenberg, B. (1958). Is the F-scale irreversible? *Journal of Abnormal and Social Psychology*, 46, 457-469.
- Eisinga R. N., Scheepers P. L. H. (1989). *Etnocentrisme in Nederland*. Nijmegen: ITS.
- Kleinpenning, G. (1993). *Structure and content of racist beliefs. An empirical study of ethnic attitudes, stereotypes and the ethnic hierarchy*. Utrecht: ISOR.
- Meloën, J. D. (1983). *De autoritaire reactie in tijden van welvaart en crisis*. Academisch proefschrift, Universiteit van Amsterdam.
- Meloën, J. D. (1991). De autoritarisme-concepten van Adorno et al. en Altemeyer vergeleken: theoretische analyse en empirische test. In P. Scheepers & R. Eisinga (Red.), *Onderdanig en intolerant. Lacunes en controverses in autoritarisme-studies*. Nijmegen: ITS.

Peabody, D. (1961). Attitude content and agreement set in scales of authoritarianism, dogmatism, anti-semitism, and economic conservatism. *Journal of Abnormal and Social Psychology*, 63(1), 1-11.

Raaijmakers Q. (1986). Verrechtsing en politieke intolerantie bij middelbare scholieren. In M. Matthijssen, W. Meeus, F. van Wel (Red.), *Beelden van Jeugd* (pp. 48-68). Groningen: Wolters Noordhof.

Raaijmakers, Q. (1993). Opvattingen over politiek en maatschappij. In W. Meeus, H. 't Hart (Red.), *Jongeren in Nederland*. Amersfoort: ACCO.

Samelson, F., & Yates, J. F. (1967). Acquiescence and the F scale. Old assumptions and new data. *Psychological Bulletin*, 68(2), 91-103.

Schuman, H., Bobo, L., & Krysan, M. (1992). Authoritarianism in the general population: The education interaction hypothesis. *Social Psychology Quarterly*, 55(4), 379-387.

Vollebergh, W., Raaijmakers, Q., & Meeus, W. (1995). Traditionele opvattingen en rechtsextremisme: trends van 1970-1994 onder jongeren en volwassenen. *Jeugd en Samenleving*, 1/2, 3-21.

Vollebergh W. (1991). *The limits of tolerance*. Academisch proefschrift, University of Utrecht.

Manuscript aanvaard 15-9-1995

Auteur

W. A. M. Vollebergh is verbonden aan de Vakgroep Jeugd, Gezin en Levensloop van de Universiteit Utrecht.

Adres: Postbus 80.140, 3508 TC Utrecht.

Abstract

Authoritarianism and ethnocentrism in pupils from lower educational levels: acquiescence or opinion?

W. A. M. Vollebergh. *Pedagogische Studiën*, 1996, 73, 123-135.

In this article it is investigated whether the high agreement with authoritarianism and ethnocentrism found in pupils in lower educational levels in the Netherlands is the result of a tendency to agree with items irrespective of item content. In the research, items expressing positive and negative attitudes towards authoritarianism and ethnocentrism were used. Results did not support the acquiescence hypothesis for the higher agreement of the lower educated pupils. Pupils from lower levels of education did show greater difficulty in expressing an opinion with respect to the items, but no relationship was found between this greater lack of opinion and agreement with items either. The strong educational effect found with respect to agreement with authoritarianism and ethnocentrism thus seems to reflect genuine differences in opinion between pupils originating from different levels of education. The nature of these opinions was highlighted by the relationship between the scores and the political party preferences of the pupils. The research was carried out in the context of the national school elections in the Netherlands, in which pupils in secondary education are asked to express their political party preferences.

Realisering Weer Samen Naar School-beleid afhankelijk van wijziging bekostigingssysteem

C.J.W. Meijer

Recent heeft Staatssecretaris Netelenbos met de Akkoordpartners (besturen-, ouder- en vakorganisaties) overeenstemming bereikt over een nieuwe bekostiging van speciale zorg. In de toekomst zal een belangrijk deel van de middelen die nu direct toevloeien naar de scholen voor speciaal onderwijs (althans voor wat betreft de zogenaamde groep-1 scholen: LOM-, MLK- en IOBK-scholen) na een geleidelijk overgangstraject worden overgeheveld naar de scholen voor basisonderwijs. De scholen voor speciaal onderwijs kunnen dan nog een genormeerd bedrag tegemoet zien op basis van een deelnamecijfer van 2% (het landelijk gemiddelde deelnamecijfer aan de groep-1 scholen is nu 3,8%). De overige middelen (op basis van de resterende 1,8%) gaan dan naar de basisscholen. Deze scholen overleggen vervolgens in het samenwerkingsverband (SWV) hoe de middelen zullen worden besteed.

Daarmee is sprake van een gedeeltelijke invoering van het oorspronkelijke streefbeeld waarin werd geanticipeerd op een volledige overheveling van zorgmiddelen. Verder is er sprake van een vereveningsprincipe. De huidige regionale verschillen in voorzieningen zullen op termijn verdwijnen. In de toekomst krijgen alle SWV's een zorgbudget op basis van het totaal aantal leerlingen. De verevening heeft een 'dubbel' karakter: zowel de financiering van speciale zorg van de in de SWV's samenwerkende basisscholen als de directe bekostiging van de speciale voorziening wordt verevend en niet meer gebaseerd op het aantal gediagnosticeerde en/of toegelaten leerlingen maar op het totaal aantal leerlingen in een gebied.

De Stuurgroep Evaluatie WSNS heeft eerder op het belang gewezen van de verandering van de bekostigingsstructuur. In het tussenrapport van de Stuurgroep (Meijer, 1995) werd betoogd dat zonder een verandering in de bekostiging er eigenlijk niet veel verwacht kan worden van de inspanningen om WSNS te realiseren. Dat de bekostiging een belangrijke voorwaarde is voor de realisering van de gewenste zorg blijkt wel uit de recente ontwikkelingen in de volksgezondheidszorg. In een recente studie die is verschenen in opdracht van de Stuurgroep Evaluatie WSNS (Meijer, Peschar & Scheerens, 1995) hebben we vastgesteld dat de problemen die gelden in de gezondheidszorg voor een deel dezelfde zijn als de WSNS-problematiek. Aldaar beproefde oplossingen kunnen dan ook worden benut bij de doordening van de condities voor de realisering van het Weer Samen Naar School-beleid. In de gezondheidszorg wordt al sinds geruime tijd expliciete aandacht geschonken aan de prikkelstructuur en de (potentiële) gevolgen van wijzigingen daarin. Bovendien zijn in de gezondheidszorg voor een deel dezelfde doelen en contextkenmerken aan de orde als bij de zorg voor leerlingen met specifieke behoeften. Ook daar voert immers het begrip 'zorg op maat' de boventoon en ook daar wordt voortdurend en intensief nagedacht over de positie van de verschillende actoren (cliënt, hulpverlener, instelling, regio en overheid) in dat kader. De studies van Hofland en Wilms (1984), Post en Vennix (1992) en het werk van de Commissie modernisering curatieve zorg (1994a en b) gaan uitgebreid in op de 'ziekte' van de gezondheidszorg en doen voorstellen voor de verbetering van de situatie. Maar ook op terreinen als de sociale werkvoorzieningen (zie de studie van BBO Consultancy (1993), de residentiële jeugdhulpverlening (we wijzen op het werk van de Commissie Harmonisatie van Normen in de periode 1988-1991), de algemene geestelijke volksgezondheidszorg (De Bruine & Van der Grinten, 1990) en de zwakzinnigenzorg (De Haen, 1992) is de discussie over een zorg op maat nauw verbonden met alternatieve bekostigingsmodellen.

Welke zijn nu die lessen uit de gezondheidszorg? We onderscheiden de volgende:

- 1) De trends in de gezondheidszorg zijn

vergelijkbaar met de ontwikkelingen in de WSNS-discussie. Het gaat in beide sectoren om de zorgbehoefte van een relatief onzekere cliënt die streeft naar vermindering van onzekerheid en verhoging van welbevinden.

2) In beide sectoren wordt verder benadrukt dat het van groot belang is dat er een continuüm van zorg ontstaat dat voldoet aan de criteria: zo kortdurend mogelijke zorg, in zo'n licht mogelijke vorm, die tijdig wordt geboden en zo nabij mogelijk.

3) De kostenbeheersingsvraag is een belangrijk punt in de gezondheidszorg. Men wil van een 'open eind' aanbod-geïnduceerde financiering naar een meer gesloten financieringsvorm waarbij de vraag meer centraal staat.

4) Er is sprake van een algemene trend tot verdergaande decentralisatie. Zorg op maat vereist lokale/regionale afwegingen die niet door de centrale overheid kunnen worden gemaakt. De overheid moet voorwaarden scheppen die een meer regionale besluitvorming stimuleren. De overheid treedt terug en zorgt voor een vermindering en vereenvoudiging van de regelgeving. Het accent komt meer te liggen op controle, informatieverstrekking, evaluatie, toezicht en kwaliteitsbewaking.

Verder wordt regionale samenwerking in de gezondheidszorg van groot belang geacht. Oplossingen voor in de regio ontstane problemen moeten ter plekke kunnen worden gegeneerd.

5) De aandacht voor de prikkelstructuur is zeer expliciet in de gezondheidszorg. De prikkels in de bekostigingsmodellen beïnvloeden de vorm, inhoud en kwaliteit van de zorg en de prijsstelling. De voorzieningen zelf beïnvloeden de vraag in vele sectoren van de gezondheidszorg. Er zijn vrijwel geen kosten/batenafwegingen in het besluitvormingsproces over het zorgtraject ingebouwd. Dit hangt samen met de sterke positie van belanghebbenden in de besluitvorming. Inmiddels wordt de financiering gaandeweg losgekoppeld van het gebruik van het zorgaanbod.

De 'incentive'-structuur moet gericht zijn op het gedrag van de belangrijkste beslissers over de zorg. Zowel aan de aanbodkant als aan de vraagkant moeten de juiste prikkels (leidend tot een economisch gebruik van kwalitatief goede zorg) worden ingevoerd.

6) Er is een verschuiving waar te nemen

van een aanbodgerichte benadering naar een meer marktgerichte (vraaggestuurde) benadering van de zorg. De cliënt krijgt een meer centrale plaats toebedeeld. Het produkt, de prestatie en de functie worden gaandeweg belangrijker parameters in bekostigingsmodellen dan de aanbodparameters. Een volledig marktmechanisme wordt niet aanbevolen in de collectieve sector waar het gaat om de zorg van een in principe in omvang beperkte risicogroep.

7) Er is maar een beperkte samenhang tussen de regionale verschillen in financiering van de zorg en de zorgbehoefte. Harmonisatie, verevening en 'vlakke' financieringsmodellen worden op verschillende terreinen ingevoerd. Behalve dat het consequenties zal hebben voor de instellingen die hun ontstaan danken aan historisch gegroeide aanbodverschillen, zijn er maar weinig argumenten tegen het doorvoeren van verevening op tal van terreinen in de gezondheidszorg. De verdeling van voorzieningen is nu veelal onrechtvaardig.

De gezondheidszorg is lange tijd beheerst door een financieringssysteem dat vanuit het aanbod (de voorzieningen) is opgezet. Dit heeft volgens velen geleid tot een enorme overconsumptie van medische produkten. Doelmatigheidsoverwegingen speelden hierbij geen rol van betekenis. Gaandeweg gaat men binnen de gezondheidszorg over tot andere financieringsmodellen. De belangrijkste zijn het zogenoemde cliëntgebonden budget en het voucher-systeem, het systeem van budgetfinanciering en het marktmechanisme.

Het cliëntgebonden budget (cgb) en het voucher-systeem vertrekken vanuit een emancipatoire gedachte: de cliënt is mondiger geworden en minder afhankelijk van het inzicht van professionals. De systemen worden gedefinieerd door het toekennen van een budget voor, respectievelijk door het recht op, een bepaalde hoeveelheid dienstverlening aan de cliënt die vervolgens zelf zijn of haar zorg 'inkoopt'. Een duidelijk probleem bij beide systemen is de mogelijk aanzuigende werking en de daaruit voortvloeiende overconsumptie. Dit doet zich vooral voor wanneer de indicatiestelling op grond van relatief 'zachte' criteria geschiedt. Dit vereist dan weer een zeer sterke protocollering van de besluitvorming en een (van het aan-

bod) onafhankelijke indicatiestelling. In het algemeen echter is de indicatiestelling een zeer problematisch onderdeel. Vooral de theoretische definiëring van de doelgroep en de praktische meetbaarheid van de kenmerken die de doelgroepen van elkaar onderscheiden leveren problemen op.

Het marktmechanisme ten slotte, waarmee cliënten, al dan niet verzekerd, zelf het gebruik van de zorg betalen, kent een aantal duidelijke gevaren. Naast mogelijke monopolisering kan een spanningsveld ontstaan tussen individuele en collectieve behoeften (met hogere premies voor de risicogroep). Het marktprincipe kan een bedreiging betekenen voor het solidariteitsprincipe. Wel kan geconstateerd worden dat het marktdenken mede heeft bijgedragen aan een heranalyse van de prikkels die een rol spelen bij de gezondheidszorg.

Het budgetfinancieringssysteem wordt op tal van terreinen binnen en buiten de gezondheidszorg toegepast, en daar waar van toepassing nog geen sprake is, is het sterk in opkomst. Het systeem omvat vele submodellen en kan bovendien op een aantal 'marktconforme' parameters, zoals de prestatie en de output, worden gebaseerd (Vrancken, 1987). Budgetfinanciering bevat drie elementen: een maximum aan toegekende middelen, de vrijheid om binnen de begroting te schuiven en de koppeling van het budget aan de prestaties of activiteiten van een instelling (Van Selm & Herweijer, 1987). Deze wijze van financiering komt steeds meer voor. Daar zijn volgens Van Selm en Herweijer drie redenen voor: er is behoefte aan kostenbeheersing, het past in het concept 'sturen op afstand', en het komt voort uit de behoefte om 'waar voor je geld' te krijgen.

Budgetfinanciering is ingevoerd voor bejaardenoorden (sinds 1985), in de intramurale gezondheidszorg (sinds 1982), het orkestwezen (sinds 1987), de nieuwbouw in de sociale sector (sinds 1987) en op tal van andere terreinen.

Vrancken (1987) definieert budgetfinanciering als: '...die vorm van subsidiëring waarbij vooraf door een budgetgever aan een budgethouder voor een bepaalde periode een maximumbedrag aan middelen ten behoeve van een zeker niveau van activiteiten en/of prestaties wordt toegekend' (p. 8).

Is zo'n budgetfinancieringsmodel nu een adequaat model voor WSNS? Vanuit het buitenland lijkt er inderdaad ondersteuning te zijn voor het (geoommerkte) budgetfinancieringsmodel. Dit model zou meer recht doen aan de behoefte om lokale/regionale oplossingen te creëren die toegesneden zijn op de betreffende situatie. Een van de alternatieven die in dit verband wel worden voorgesteld is het leerlinggebonden budget (of ook wel het rugzak-model genoemd). Vanuit onze studie naar dergelijke modellen in de gezondheidszorg wezen we reeds op een aantal duidelijke nadelen. Buitenlandse gegevens met betrekking tot soortgelijke modellen, maar dan toegepast in het onderwijs bekrachtigen dit: het aan leerlingen gekoppelde budget gaat veelal gepaard met bureaucratische diagnostische procedures (vaststelling van de indicatie) en heeft een aanzuiende werking.

In ons land is op dit moment nog sprake van een aanbodgerichte bekostiging van de speciale zorg. Onderzoekers wijzen op de rol die het aanbod in ons land speelt. Zo zijn er grote regionale verschillen in deelname aan het speciaal onderwijs geconstateerd. Er zijn indicaties dat die verschillen vooral samenhangen met regionale verschillen in het voorzieningenniveau. Diezelfde onderzoekers pleiten dan ook voor een verevening van de middelen die voor de speciale zorg voor leerlingen beschikbaar zijn.

Reeds eerder (zie *Halverwege*) wezen we erop dat zolang de middelenstructuur niet verandert, er in feite niet veel te verwachten is van de WSNS-inspanningen. Immers, waar aan de ene kant 'zorgverbreding en differentiatie worden nagestreefd is er aan de andere kant ook sprake van 'incentives' in de regelgeving die de groei van het SO bevorderen. In de praktijk betekent dit dat het toelaten van leerlingen in het SO wordt beloond, het terugplaatsen niet (voldoende) wordt bevorderd en ook het handhaven van leerlingen met problemen in het reguliere stelsel niet wordt gestimuleerd. Het vigerende 'tweesporen'-systeem belooft dus segregatie en ontmoedigt integratie. De ervaringen met het WSNS-proces wijzen erop dat veel scholen strategisch gedrag vertonen. De wijze van bekostiging draagt mede bij aan het feit dat scholen steeds weer nieuwe wegen zoe-

ken om de financiën op peil te houden.

Natuurlijk doemt hier de vraag op of de prikkelstructuur in het door de Akkoordpartners ontwikkelde model meer gericht is op het realiseren van de WSNS-doelen dan de vigerende bekostigingsstructuur. Dit lijkt inderdaad het geval te zijn: zowel de verevening, waar in de voorstellen sprake van is, als de (gedeeltelijke) overheveling van zorgmiddelen sporen meer met de doelstellingen van het WSNS-beleid dan een niet-verevende directe bekostiging van speciale scholen. In die zin is sprake van een beter met de doelen in overeenstemming zijnde 'incentive'-structuur. Meer verwijzen wordt in dat model immers niet meer beloond en de opvang op reguliere scholen niet meer ontmoedigd. Daarmee wordt de kans dat onderwijskundige overwegingen een dominante rol spelen bij beslissingen over het gewenste zorgaanbod en de plaats waar dat dient te geschieden gemaximaliseerd. De bekostiging is dan min of meer 'plaatsing-neutraal' geworden en is daarmee in overeenstemming met de aanbevelingen die we uit buitenlandse gegevens en uit andere sectoren zoals de gezondheidszorg hebben kunnen vernemen.

Literatuur

- BBO Consultancy (1993). *Onderzoek naar een financieringsmethode voor de SW.* 's-Gravenhage: Ministerie van Sociale zaken en Werkgelegenheid.
- Bruine, M. de & Grinten, T. E. D. van der (1990). *Sturing van geestelijke gezondheidszorg door overeenkomsten. Een economische analyse.* Rotterdam: EUR.
- Commissie Harmonisatie van Normen (1988). *Harmonisatie van normen op het terrein van de jeugdhulpverlening. Advies aan de minister van WVC en de staatssecretaris van Justitie. Deel 1: Hoofdpijnen.* 's-Gravenhage: DOP.
- Commissie Harmonisatie van Normen (1991). *Harmonisatie van normen op het terrein van de jeugdhulpverlening. Advies aan de minister van WVC en de staatssecretaris van Justitie. Deel 3: Eindadvies.* 's-Gravenhage: DOP.
- Commissie modernisering curatieve zorg (1994a). *Achtergrondstudies 'Gedeelde zorg: betere zorg'* (achtergrondstudies van het NIVEL en het NZi). Zoetermeer: Hageman b.v.

- Commissie modernisering curatieve zorg (1994b). *Gedeelde zorg: betere zorg.* Zoetermeer: Hageman b.v.
- Haen, F. de (1992). *Bekostiging op niveau?* Utrecht: Nationaal Ziekenhuisinstituut.
- Hofland, J., & Wilms, P. J. M. (1984). *Onder behandeling: strategische keuzen bij kostenbeheersing in de gezondheidszorg.* Den Haag: Instituut voor onderzoek van overheidsuitgaven.
- Meijer, C. J. W. (1995). *Halverwege: van Streefbeeld naar Startwet.* De Lier: Academisch Boeken Centrum.
- Meijer, C. J. W., Peschar, J. L., & Scheerens, J. (1995). *Prikkels.* De Lier: Academisch Boeken Centrum.
- Post, D., & Vennix, J. A. M. (Red.). (1992). *De gezondheidszorg in model. Ons complex zorgsysteem gemeten en gewogen.* Houten/Zaventem: Bohn Stafleu Van Loghum.
- Selm, J. A. van, & Herweijer, M. (Red.). (1987). *Budgetfinanciering: een beleidsverkenning.* Groningen: Samenwerkingsverband Bestuurswetenschappen Noorden des Lands.
- Vrancken, P. H. J. (1987). Budgetfinanciering in de collectieve sector. In J. A. van Selm & M. Herweijer (Red.), *Budgetfinanciering: een beleidsverkenning* (pp. 7-18). Groningen: Samenwerkingsverband Bestuurswetenschappen Noorden des Lands.

Weer Samen Naar School: als het doel op is, beginnen we aan de middelen

Reactie op de bijdrage van

C.J.W. Meijer

A.J.J.M. Ruijsseenaars

Weer Samen Naar School heeft als beweging een goed gekozen 'naam' meegekregen. De Verloren Zoon had qua betekenis ook gekund, maar heeft niet helemaal dat positieve en toekomstgerichte. Wat het weer wel heeft, is dat de prikkelstructuur voor zoonlief van het juiste type was: geld blijft trekken.

Speciale zorg en geldstroom

'Recent heeft Staatssecretaris Netelenbos met de Akkoordpartners (besturen-, ouder- en vakorganisaties) overeenstemming bereikt over een nieuwe bekostiging van speciale zorg.' Prachtig. En het klinkt veel opgewekter dan: 'Men probeert de vraag naar hulpverlening te verminderen door de geldstroom te beteugelen.' Deze laatste zin werd opgetekend uit de mond van Prof.dr. A. van Dantzig (De Volkskrant, 28 oktober 1995) naar aanleiding van diens opvattingen over het veel te rooskleurige beeld van het gezin. De titel van het interview-artikel luidde overigens veelbetekend: 'Er is veel meer gekte dan je denkt.' Van Dantzig wijst er op dat het verminderen van de geldstroom ('een nieuwe bekostiging') als middel om de vraag naar hulpverlening ('speciale zorg') terug te dringen, leidt tot het creatief zoeken en vinden van hulp op andere plaatsen. 'Dan krijg je wachtlijsten, mensen die afhaken, naar het alternatieve circuit gaan. En vooral: de neiging bij de rijken om hun voorzieningen te kopen en hun heul te zoeken bij privé-klinieken, privé-dokters, privé-therapeuten buiten de collectieve en verzekerde voorzieningen om. Die tweedeling is nu aan het ont-

staan.' Van Dantzig bedoelt overigens niet dat de rijken ongelijk hebben en lijden aan luxe-problemen, maar dat de (ook speciale) voorzieningen voor geestelijke gezondheidszorg voor iedereen bereikbaar moeten zijn en blijven. Er wordt niet voor de lol om speciale zorg gevraagd. Ook niet in het (speciaal) onderwijs.

In Nederland krijgen relatief veel leerlingen speciale zorg in het speciaal onderwijs. Ouders en leerlingen blijken door de bank genomen tevreden te zijn over deze basisvoorziening, ook al mag de taxi best wat later komen en weer wat eerder terug zijn, en kan enig pedagogisch inzicht bij de chauffeur geen kwaad. Toch lijkt het er op dat we ons voor deze vorm van special zorg dienen te schamen. We segregeren en stigmatiseren dat het een lieve lust is, en het kost nog handen vol geld ook. Nee, kijk dan naar andere landen. Die doppen hun boontjes geïntegreerd of slagen er in de groei naar hulpverlening op z'n minst te stabiliseren. Het beeld dat daarvan wordt geschetst zou wat mij betreft een onafhankelijke kritische analyse waard zijn. Enige reserve blijkt dan wellicht niet misplaatst. Neem ik de stabiele omvang van het speciaal onderwijs in Vlaanderen als voorbeeld (nog even los van de scholenstop voor het speciaal onderwijs tussen 1971 en 1986), omdat ik die situatie van dichtbij enigszins ken. Het totale deelnamepercentage is inderdaad stabiel en blijft al lange tijd steken rond de drie procent. Vlaanderen kent echter geen of nauwelijks speciale voorzieningen voor kleuters en voor leerlingen van het voortgezet onderwijs, twee bronnen van sterke groei in Nederland. Bovendien zijn de percentages binnen het speciaal onderwijs zelf verre van stabiel: een forse afname van het aantal leerlingen in type-1 scholen (MLK) en een grote toename van leerlingen in type-8 scholen (LOM). En dit laatste ondanks het feit dat het gemiddeld aantal leerlingen in het gewone onderwijs per klas laag is (dikwijls rond de twintig) en scholen kunnen beschikken over een taakleerkracht en een zogenaamde aanpassingsklas. Buiten het onderwijs zien we dan nog een sterke groei in privé-hulpverlening en bestaat er een dicht landelijk netwerk van ambulante revalidatiecentra (gezondheidszorg) waar aan kinderen met leerproblemen hulp wordt gegeven door vooral logopedisten en motorisch-therapeuten, in mindere

mate ook door gedragswetenschappers.

Stabiliteit is kennelijk relatief, maar groei en afname zijn dat natuurlijk ook. Beleidsmatig zijn ze als verschijnsel eenvoudig hanteerbaar. Als je geen absolute groei meer wilt van het speciaal onderwijs van enig type, dan is een maatregel in de vorm van een klassen- en scholenstop aan te bevelen, of anders de afspraak dat er voor elke uitstromer slechts één instromer mag zijn. Wie mag instromen, moet door loting eenvoudig te bepalen zijn, bijvoorbeeld door een regionale lotingscommissie. Wil je een bepaalde afname van het absolute aantal leerlingen, dan is het redelijk om bijvoorbeeld voor elke vier uitstromers er slechts weer drie toe te laten. Je houdt dan op den duur weliswaar leerkrachten over, maar die laat je gedeeltelijk in het speciaal onderwijs meer expertise opbouwen en kun je tegelijk in het basisonderwijs inzetten als extra steun voor die uitgelote pechvogel. Dat voorkomt ideeën als het aanstellen van klasseassistenten, waardoor er binnenkort twee mensen per lokaal met hun handen in het haar zitten. Toegegeven, scheefgegroeide regionale verschillen in voorzieningen blijven daarmee bestaan, maar dat is alleen een probleem als je de hele samenleving wilt regionaliseren en er geen sturende landelijke verantwoordelijkheid meer mag zijn of als het niet lukt om bijvoorbeeld scheidslijnen tussen onderwijs en gezondheidszorg op te lossen.

Geen geld, geen integratie?

De voorgaande gedachtengang is ongetwijfeld te zwart-wit, maar in het idee dat je voor de speciale hulp aan leerlingen met leerproblemen de middelen onttrekt aan het speciaal onderwijs en aan de daar opgebouwde expertise kan ik ook nauwelijks grijs tinten ontdekken. Ik ken het tegenargument: het geld wordt niet zozeer weggehaald (het is geen bezuiniging), maar het wordt anders gebruikt. Dat veronderstelt natuurlijk dat het niet uitmaakt hoe je het geld uitgeeft. Alsof je mag kiezen tussen drie verpleegsters of één hartspecialist. Een ander tegenargument is dat je de problemen moet aanpakken waar ze ontstaan: in de interactie tussen leerling en leerkracht. De foutieve gevolgtrekking die velen daaraan verbinden is

echter dat dit ook betekent dat leerproblemen ontstaan *door* die interactie, een conclusie waartegen Dumont al jaren geleden waarschuwde.

Een voordeel van de hierboven gegeven administratieve (en goedkope!) zwart-wit-oplossing is dat je geen omslachtige filosofieën nodig hebt over bijvoorbeeld het mooie van integratie. Want wie is daar nou tegen? De leerkracht in het speciaal onderwijs? De leerkracht in het basisonderwijs die na veel inspanning geen weg meer ziet? De ouders die hulp zoeken? De steun voor de zwart-wit benadering wordt door Meijer trouwens goed verwoord: 'In het tussenrapport van de stuurgroep (*Halverwege*) werd betoogd dat zonder een verandering in de bekostiging er eigenlijk niet veel verwacht kan worden van de inspanningen om WSNS te realiseren.' Maar waarom zou iets dat intrinsiek goed is, dat beleidsmatig stevig ondersteund wordt en dat tot een gelukkiger samenleving zal leiden niet werken? Het antwoord dat de juiste prikkelfstructuur ontbreekt, gaat me net even een stap te snel. Het punt is immers niet waar je het geld heen stuurt, maar waar de expertise zit en waar je het geld het beste kunt investeren. Nogmaals: drie verpleegsters kunnen niet gemist worden, maar je hebt nu eenmaal ook specialisten nodig. Leerproblemen van leerlingen in het speciaal onderwijs zijn doorgaans uiterst hardnekkig en gaan niet zelden gepaard met een aantal andere moeilijkheden. Inmiddels wordt steeds vaker gesproken van meervoudige problemen, die niet alleen vragen om gedifferentieerd werken en een goed systeem van klasmanagement, maar ook om tijd-rovende en op het individu toegesneden hulp, met steun vanuit verschillende deskundigheden en met ondersteuning voor de betrokken opvoeders. En nogmaals: kinderen zitten niet voor hun lol in het speciaal onderwijs (ook al komt bij een redelijk aantal het plezier in school weer terug), zoals ook ouders niet vanzelfsprekend een plaatsing vragen, en ook de verwijzende leerkracht niet uit gemakzucht handelt of de verantwoordelijkheid liever kwijt dan rijk is.

Doel en middelen

Budgetfinanciering en prikkelstructuur: het doel raakt uit beeld en de middelen gaan de discussie bepalen. De ontwikkelingen in de gezondheidszorg zetten de trend. Meijer: 'De gezondheidszorg is lange tijd beheerst door een financieringssysteem dat vanuit het aanbod (de voorzieningen) is opgezet. Dit heeft volgens velen geleid tot een enorme overconsumptie van medische producten. Doelmatigheidsoverwegingen speelden hierbij geen rol van betekenis.' Dit spreekt tot de verbeelding, zij het dat Minister Borst en Staatssecretaris Terpstra in de nota 'Prijzbeleid en financieel beleid zorgsector' eind oktober wezen op het nog altijd uit de hand lopen van de kosten. Ook de bejaardenoorden kennen het nieuwe financieringssysteem (sinds 1985). Een directeur die ik daarover opbelde, klonk redelijk tevreden: 'Je zorgt ervoor dat de bezettingsgraad boven de 99,5% blijft (volle bak-principe) en je hebt nu de vrijheid om bepaalde mensen niet op te nemen of om bepaalde (moeilijk definieerbare) zorg niet te geven. Maar je weet vooraf wat je krijgt.' Het orkestwezen blijft sinds 1987 evenmin achter. Een telefoontje naar een van onze grootste orkesten leverde evenmin gemopper op: 'De vrijheid van handelen is prima en met sponsoring lukt het goed. Problemen ontstaan natuurlijk wel als het geld per regio gaat worden toegewezen en het geld over verschillende zalen (wij spreken in zalen) moet worden verdeeld. Maar ja, wij zijn een groot orkest, dus dan hebben die kleine zalen in elk geval geld om ons te kunnen laten spelen.' Kortom: hoe beheersbaar worden de kosten, wie bepaalt welke zorg wordt gegeven, hoe komen ook diegenen aan de bak die wat minder geluid maken en wat gebeurt er met de kleine 'zalen'?

We zijn er niet klaar voor, maar gaan wel alvast

De Staatssecretaris en de Akkoordpartners zijn het dus eens. Toch maar even gebeld met een niet onaanzienlijke landelijke vereniging van ouders met kinderen die doorgaans het speciaal onderwijs bezoeken ('We zijn te laat op de kar gesprongen en zijn dus geen officiële akkoordpartner'): 'Met ons is niet structureel overlegd

en we hebben niet kunnen tekenen. Over het algemeen blijken ouders niet of erg slecht geïnformeerd. De voorgestelde regeling werkt wel voor een minimumpakket aan voorzieningen, maar ouders die meer hulp willen, zullen het zelf gaan bekostigen. Bovendien is de kans groot dat bij de regionale verdeling van middelen (en door het ontbreken van landelijke richtlijnen) de grootste schreeuwers winnen en de pot snel leeg is. Maar het grootste probleem is wel dat het basisonderwijs er absoluut nog niet klaar voor is.'

Tot slot

Het aanpakken van flinke problemen vraagt meer dan pedagogisch optimisme en onderwijskundige differentiatie, hoe veel bewondering ik ook heb voor de inzet van onderwijsgevers en begeleiders die hard werken aan WSNS. Dat 'meer' is in de vorm van expertise in het speciaal onderwijs aanwezig en vraagt om een blijvende investering. In plaats van het uitbouwen en 'exporteren' van die deskundigheid kiest het beleid voor het uitsmeren van middelen onder de noemer van integratie. Inmiddels blijken de middelen het van het nobele doel te gaan winnen. Wie weet wanneer de verloren zoon weer thuis komt.

Discussie

Prikkels tot zorgzaam onderwijs?

Reactie op de bijdrage van

C. J. W. Meijer

G. H. van Gemert

Inleiding

Kinderen gaan blijkbaar niet vanzelf weer samen naar school. Dat is logisch, vindt Meijer, want de bestaande vormen van financiering bevorderen bepaald niet de zo wenselijke 'korte, lichte, tijdige en nabije' zorg. Dus acht men 'prikkels' nodig, bestuursaal voor 'financieel afdwingen'. Men denkt zo op macro-niveau het belang van leerlingen te dienen. Meijer concludeert dat de komende maatregelen rond budgettering en verevening 'plaatsing-neutraal' zijn, en wijst op voor- en nadelen van de persoonsgebonden budgetten, zoals die af te leiden zouden zijn uit de ervaringen in binnen- en buitenland op het gebied van de gezondheidszorg. Hij ziet al met al vooral de zonzijde van de voorstellen. Dat vraagt om een reactie.

Financiële prikkels

Voortdurend duikt in het spreken en schrijven over WSNS het woord 'kostenbeheersing' in combinatie met 'overconsumptie' op (ook in Meijers stuk), wat doet vermoeden dat veranderingen moeten worden doorgevoerd onder de hypotheek van voorgenomen bezuinigingen. 'Weer' en 'samen' zijn dan gewoon nieuwspraak voor reductie van kostbare vormen van zorg. Wie stelt eigenlijk vast wat 'teveel' is, op welke gronden? Welk maximumbedrag wordt vastgesteld bij budgettering en verevening door wie? Verevening = ieder even weinig? Hoe opportunistisch (politiek getint) is die discussie? Maar zelfs met voorbijgaan aan die achtergrond is het de vraag of Meijers voorzichtig positieve oordeel over de komende fi-

nancieringsstructuur gerechtvaardigd is. Ik betwijfel of financiële 'incentives' inderdaad wel in de goede richting sturen. 'Middelen' zijn natuurlijk niet weg te denken, maar behoren het sluitstuk van een doelgerichte redenering te zijn. Zoals het pakket maatregelen nu gepresenteerd wordt, doet het vooral een beroep op de calculerende schoolleider om niet te lichtvaardig door te verwijzen. Een kwestie van zwaardere argumenten zoeken vermoedelijk.

Dat komt bekend voor: ook in de gezondheidszorg is al vaak geprobeerd de kosten te drukken door de middelen te beheersen, met betrekkelijk geringe resultaten, behalve dat de dagelijkse zorgverleners steeds zwaarder belast werden. Misschien moeten we toch nog iets nauwkeuriger kijken naar de voorbeelden die Meijer noemt. Er zijn namelijk nog wel meer lessen te trekken uit de ontwikkelingen in de gezondheidszorg die hij noemt. Ik wil er slechts twee noemen, want dit moet een korte reactie worden. De belangrijkste is dat de bestaande infrastructuur (en nu bedoel ik meer dan alleen de financiële middelen) niet zomaar bestaat, maar berust op ideeën. Mensen legitimeren hun handelen door naar die ideeën te verwijzen. Zolang de laatste niet veranderen, zullen nieuwe systemen worden uitgehold of geperverteerd, net zolang tot ze onschadelijk zijn of het oude systeem zelfs versterken. Veranderingen zonder beïnvloeding van de ideeënwereld van betrokkenen zijn tot mislukken gedoemd. En de tweede belangrijke les is dat persoonsgebonden budgettering geen garantie vormt voor individuele aandacht, zolang het niet wordt opgehangen aan planning met aandacht voor het eigene van het kind. Hetgeen overigens iets anders betekent dan dat alle zorg of onderwijs individueel verleend moet worden. Niet ieder kind vraagt om een individueel traject, ook al denken de opvoeders kindgericht. Deze opmerkingen licht ik in het navolgende toe.

Weer samen naar school?

Bladeren in beleidsstukken en artikelen over WSNS levert voornamelijk mist op, bestaande uit schijnbaar vanzelfsprekende basis-ideeën, vaagheid over de onderbouwing van die ideeën en over de concrete gang van zaken op micro-

niveau in de scholen zelf. Slechts zelden wordt bijvoorbeeld onderscheid gemaakt tussen normalisatie en integratie. Het verschil is natuurlijk wie zich aanpast: respectievelijk het individu of de maatschappij. WSNS mikt officieel op integratie, probeert marginalisering van moeilijke leerlingen te voorkomen. Maar veel ruimte voor het van de maatschappelijke norm afwijkende kind wordt niet geboden. De woorden 'weer' en 'samen' duiden op een romantische ideologie waarin alle kinderen zonder onderscheid, net als vroeger, naar de basisschool gaan. Of die goede, oude tijd ooit bestaan heeft, laten we maar in het midden. Maar hoe het dan in onze tijd moet met de aanpak van lastige, domme of onwillige leerlingen (of leerlingen met een combinatie van deze eigenschappen), staat nergens. WSNS als 'back to basics'. Als dat maar geen terugval wordt. Ter vergelijking: in de zorg voor gehandicapten blijkt het heel lastig om in slechts fysiek geïntegreerde (lees: genormaliseerde) voorzieningen nog iets van individuele, specialistische zorg te realiseren, bijvoorbeeld in het geval van ernstige gedragsproblemen. Het normale is vaak erg sterk de norm. Zou dat in basisscholen anders gaan?

Een tweede punt is dat het in de stukken vrijwel altijd de leerlingen zijn die problemen hebben. Dat moet een misverstand zijn. Probleemsituaties in de gezondheidszorg zijn niet te begrijpen, laat staan te beïnvloeden zonder allerlei mensen op allerlei niveaus in de beschouwing te betrekken. Diegenen die dagelijks omgaan met de 'zorgvrager' blijken vaak de problemen in stand te houden of te intensiveren. Daar komt nog bij dat die personen, net als in het onderwijs, als professional bij de problemen zijn betrokken, met andere woorden, in probleemsituaties systematisch de verkeerde kant op werken. In het onderwijs zal ook wel waar zijn dat het reduceren van probleemsituaties tot leerlingproblemen opvoedingsproblemen vaak intact laat of verergert. Weer samen naar school vereist een verandering in grondhouding van veel onderwijsgeevenden en hun managers. Die voorwaarde wordt maar magertjes geoperationaliseerd.

Ten derde noemt men vaak (Meijer doet dat ook) emancipatie als doel van WSNS. In de gezondheidszorg krijgt dat vorm via persoonsge-

bonden budgettering. Het lijkt erop dat ook in het onderwijs zo'n constructie gaat komen (het rugzakjes-model). Maar ook dan is het maar de vraag in hoeverre de leerling of zijn ouders iets te zeggen krijgt over wat er in die rugzak zit en wat hij ervoor kan kopen. De ontwikkelingen in de gezondheidszorg stemmen niet hoopvol. Uit de gang van zaken op het gebied van de zorgverlening in Nederland kan men vooral leren dat bureaucratisering en gebrekkige indicatiestelling geen voorbijgaande uitvoeringsproblemen zijn, maar altijd daar opdoemen waar de belangen van gevestigde instanties in het geding komen. Bureaucratisering, bescherming van belangen van anderen dan de cliënt en onpersoonlijke regelingen kenmerken bijvoorbeeld de invoering van persoonlijke budgetten in de zorg voor gehandicapten. In dit verband moet de waarde van overlegorganen en indicatie-commissies voor de emancipatie van de zorgvrager niet hoog worden ingeschat. In de gehandicaptenzorg overheerst de voorziening-gerichte intake, ook daar waar de cliënt officieel centraal staat. Het samenwerkingsverband per regio dat nu wordt voorgesteld, emancipeert misschien wel de basisscholen ten opzichte van het SO, maar dat is nog niet hetzelfde als het belang van de leerling dienen. Consensus en gelijkgerichtheid zijn kwetsbaar. Een spelbreker is genoeg om de belangenstrijd per regio te doen oplaaien.

Wat stuurt de voortdurende uitstroom van de basisschool? Spreken over de 'zuigkracht' van het SO is geen verklaring, maar een beschrijving van wat gebeurt. De neiging tot 'overplaatsen' van moeilijke leerlingen is geen teken van gemakzucht, maar van pedagogisch-didactische onmacht. Onmacht te verstaan als: een combinatie van gebrek aan deskundigheid en het onvermogen om zich te onttrekken aan een schoolsysteem dat modale leerlingen goed kan bedienen maar voor afwijkingen weinig plaats biedt. Dat lijkt op instellingen voor gezondheidszorg waar normalisatie en aanpassing door het individu voorwaarde is voor handhaving van de zorgvrager. Roepen dat de zorgbreedte moet toenemen, verandert niets aan het denken over modaal en afwijkend, noch aan de prioriteitsstelling tussen beide. WSNS is een zware opgave omdat de cultuur in scholen moet worden geattaqueerd, niet omdat het SO

teveel geld krijgt. De 'prikkelstructuur' zal vermoedelijk vooral de creativiteit van schoolleiders prikkelen om hun werkbudget via 'overleg' te verhogen.

Een voorbeeld uit de gehandicaptenzorg

Er valt een duidelijke parallel te trekken tussen de benadering van 'moeilijke' leerlingen en de manier waarop in de gehandicaptenzorg in het verleden gereageerd is op het verschijnsel 'ernstige gedragsproblemen' bij mensen met een handicap. Het bijeenplaatsen van moeilijke mensen in categorale voorzieningen bleek ook daar een dure, gedeeltelijke oplossing (vergeleek de bezwaren tegen niet-geïntegreerd SO). Daarnaast heeft men lang de opvatting gehuldigd dat de instellingen voor zorgverlening zelf de oplossingen voor deze problemen moesten en konden realiseren, gebruikmakend van de bestaande mogelijkheden en middelen. Ernstige (gedrags)problemen laten zich echter niet met gebruikelijke middelen oplossen binnen het bestaande systeem, integendeel. Het failliet van deze benadering is gemarkeerd door de foto van Jolanda Venema.

In reactie op deze roemruchte casus is onder andere een systeem van consultatie in het leven geroepen. Het houdt in dat hulpverleners advies krijgen van collega's over de richting waarin oplossingen moeten worden gezocht. Het doel van de consulent is het vergroten van de deskundigheid en het handelingsrepertoire van zijn collega-zorgverlener. Hij zal dus zeer terughoudend zijn in het 'overnemen' van het probleem. Soms is de consulent meer gespecialiseerd dan degene die consult vraagt (bijvoorbeeld in diagnostiek of specifieke behandelingsmethoden), maar vaker is niet zozeer de deskundigheid van de consulent als wel zijn buitenstaander-zijn van doorslaggevende betekenis. De probleemsituatie wordt letterlijk met andere ogen gezien. Dat blijkt mogelijkheden te openen voor nieuwe, vaak succesvolle probleemoplossingen, die overigens van allerlei niveau kunnen zijn: variërend van specifieke therapie tot organisatieverandering. De consulent, een betrokken buitenstaander, brengt vaak een nieuw perspectief.

Van belang is de autonomie van de consulent. Het geld voor consultatie en de interventies die daarvan het gevolg zijn, moet dus niet door de instellingen zelf beheerd, laat staan verdeeld worden. Er is een aparte (overigens betrekkelijk smalle) geldstroom voor geschapen. De werkwijze van de consulenten wordt regionaal gecoördineerd en gefinancierd door (eveneens autonome) coördinatoren en op gezette tijden geëvalueerd. Financiering vindt plaats op grond van een op het individu gericht plan. Er is dus een persoonsgebonden aanpak van de problematiek. De coördinatoren staan onder toezicht van een bestuur dat de voorzieningen in de regio representeert, zodat er indirect een relatie met het zorgveld is. Consultaties en bijzondere zorgplannen houden zo hun tijdelijke en individu-gerichte karakter en men voorkomt dat het geld dat ermee gemoeid is, structureel deel gaat uitmaken van instellingsbudgetten. Het is, met andere woorden, geen open-eind-regeling.

Is het een idee het geld dat nu nog naar het speciaal onderwijs gaat, deels te besteden aan het opzetten van een systeem van consultatie- met-eigen-financiering? Voorwaarde is natuurlijk dat scholen en onderwijsgevers een deel van hun autonomie opgeven. Een consulent is immers meer dan een adviseur, zij het dat hij ook geen dienstbevelen kan geven. Consultatie lijkt in ieder geval een goede mogelijkheid om de discussie te verleggen naar het onderwerp waar het over behoort te gaan: het verbeteren van probleemsituaties rond moeilijke leerlingen.

Ook als men zo'n systeem in het onderwijs niet kan of wil invoeren, laat het voorbeeld van de ernstige gedragproblemen bij mensen met een handicap duidelijk zien dat een aanpak van ingewikkelde problemen om een middenweg vraagt. Het uit handen nemen van problemen, maar ook strikt intern oplossen ervan leidt tot onvoldoende resultaat. Er is een aparte inzet nodig, ook bij zogenaamd individuele problemen, om de instandhoudende cultuur te doorbreken en nieuwe perspectieven te wijzen. De financiële ondersteuning van deze interventies behoort niet in handen te zijn van belanghebbenden bij het bestaande systeem, hoe integer die ook hun eigen doelen trachten te realiseren.

Reactie op de bijdragen van A.J.J.M. Ruijsenaars en G.H. van Gemert

C.J.W. Meijer

De financiering van speciale zorg

Het openingsartikel in deze discussie is gericht op de analyse van één van de voorwaarden voor de realisatie van WSNS: de financiering van speciale zorg. Een adequate bekostiging is een noodzakelijke maar nog niet een voldoende voorwaarde voor het scheppen van de gewenste speciale zorg. Daar is vanzelfsprekend veel meer voor nodig: juiste attitude, opleiding, materialen, flexibel onderwijsleerproces, et cetera. Aan deze en andere voorwaarden heb ik in mijn bijdrage geen aandacht besteed; ik heb me beperkt tot de bekostigingsvraag.

De vigerende bekostiging is een vorm van aanbodfinanciering (van de speciale school) waarbij het aantal leerlingen dat de speciale school bezoekt de financieringsbasis is. Ik heb betoogd dat een dergelijke vorm van aanbodfinanciering (op basis van de inputparameter: aantal geplaatste leerlingen) niet zal kunnen leiden tot het bereiken van één van de kerndoelen van WSNS: het creëren van een continuüm van zorg waarbij het accent ligt op opvang binnen het reguliere stelsel. Elk alternatief financieringsmodel kan vervolgens worden beoordeeld op de mate waarin deze de kansen op het bereiken van die doelstelling maximaliseert. Het alternatief dat ik schets is het budgetfinancieringsmodel. Dit betekent dat elk samenwerkingsverband een 'verevend' budget (onafhankelijk van het aantal speciale leerlingen, maar gebaseerd op het totaal aantal leerlingen in het samenwerkingsverband) krijgt toegekend.

De reacties van Ruijsenaars en Van Gemert betreffen niet zozeer mijn betoog maar zijn hoofdzakelijk gericht op de WSNS-probleem-

stelling in het algemeen en het WSNS-beleid in het bijzonder inclusief de doelen en middelen die daarbij aan de orde zijn. Van Gemert gaat ook uitgebreid in op de voors en tegens van het leerlinggebonden budget, een financieringsmodel dat ik niet voorsta en overigens ook niet aan de orde is bij het WSNS-debat (dat slechts betrekking heeft op de groep-1 scholen: het LOM-, MLK en IOBK-onderwijs). Dit leerlinggebonden budget duikt wel op in verband met het beleid ten aanzien van de overige scholen voor speciaal onderwijs: de groep-2 en -3 scholen. De opmerkingen dienaangaande kunnen dus beter worden gericht aan de commissie-Rispens, die over die problematiek recent een advies heeft laten verschijnen. Overigens was reeds duidelijk dat staatssecretaris Netelenbos een leerlinggebonden budget voor de groep-1 scholen niet erg geschikt achtte, precies op grond van de argumenten die ook Van Gemert heeft aangedragen.

Beide reacties bevatten dus maar in beperkte mate een analyse van mijn bijdrage en gaan niet in op de vragen die mij zouden interesseren:

1. Is een vergelijking tussen WSNS en de gezondheidszorg valide?
2. Welke lessen zijn uit de gezondheidszorg te trekken met betrekking tot de financiering van zorg? Zijn dat andere dan die ik heb getrokken?
3. Welke consequenties heeft dat voor de bekostigingsvraag in het WSNS-debat? Verschillen we daarover van mening?

Het feit dat de beide reacties die vragen niet beantwoorden, zou kunnen doen vermoeden dat men weinig heeft in te brengen tegen mijn betoog. Daar kan ik natuurlijk vrede mee hebben. Afgaande op hun reacties is het echter waarschijnlijker dat beide auteurs zo vol zitten met andere vragen en problemen aangaande het beleid ten aanzien van het speciaal onderwijs in het algemeen en het WSNS-beleid in het bijzonder, dat zij amper toekomen aan de hiervoor genoemde vragen. Zij houden zich onder meer bezig met vragen als: Is integratie een goed streven? Is WSNS wel goed in de steigers gezet? Is de filosofie erachter wel goed uitgewerkt? Gaat WSNS wel lukken? Is WSNS geen ordinaire bezuiniging? Is het wel zo verstandig om het stelsel van speciale scholen uit te roeien? Is het leerlinggebonden budget wel een

goed idee? En de invoering van klasse-assistenten?

Het zijn vragen die ik hier niet heb aangevoerd en het zijn ook geen vragen die in het kader van deze polemiek aan mij kunnen zijn gericht. Mijn opvattingen over een aantal van die kwesties zijn echter duidelijk. Ik verwijs daarvoor naar de rapporten die in het kader van de evaluatie WSNS (mede) van mijn hand zijn verschenen.

Drie verpleegsters of één hart-specialist?

Het dilemma zoals in deze aanhef verwoord kan goed als leidraad dienen voor de discussie over de doelen en middelen in het kader van het WSNS-beleid. Het keuzeprobleem is verwoord in de reactie van Ruijsenaars. Met een beetje fantasie kan men voor de drie verpleegsters ook lezen: leerkrachten en ondersteunende hulp in het basisonderwijs, en kan men één hartspecialist substitueren door de deskundige in de speciale school. Dit dilemma raakt inderdaad een aantal kernvragen die voor WSNS onverkort gelden:

1. Wie maakt hier de uiteindelijke keus over de gewenste opvang?
2. Op basis waarvan wordt die keus gemaakt?
3. Zijn er nog meer opties? Zo, ja welke?
4. Wat weten we van de voors en tegens van de verschillende opties?
5. Welke procedures volgen we in het keuzeproces en bij de verantwoording?
6. Is het aan de overheid om een en ander uitputtend te regelen?

Ik beantwoord deze vragen hier niet in extenso maar geef kort een aantal overwegingen. Om met de laatste vraag te beginnen: het is niet de taak van de overheid om alle 'ins' en 'outs' aangaande de vorm en inhoud van zorgvormen (en dus ook niet van speciaal onderwijs) te regelen. Veel ervan is voorbehouden aan andere partijen. Het is zelfs van belang om niet alles uitputtend van bovenaf te regelen zodat lokaal-regionale toesnijding mogelijk wordt, rekening houdende met de verschillen in historie, context en dergelijke.

Wat betreft de keus tussen de drie verpleegsters en de ene hartspecialist: sommigen (met name

diegenen met geen, andere of lichte klachten) zullen direct hun toevlucht nemen tot het eerste alternatief, anderen (met name diegenen met ernstige hartklachten) zullen vermoedelijk voor de tweede optie kiezen. Essentieel is de vraag of deze dichotomie voldoende recht doet aan de ontwikkelingen die in de praktijk kunnen worden waargenomen. Als het gaat om het bieden van speciale zorg in het onderwijs zijn er al jaren ontwikkelingen gaande die te interpreteren zijn als een streven naar het oprichten van een continuüm van zorg. Echter de vigerende regelgeving verhinderde dergelijke innovaties. Ouders met een gehandicapt kind die er naar streven om hun kind te laten opgroeien binnen de muren van het reguliere stelsel moeten zich in bochten wringen om de daarvoor noodzakelijke faciliteiten te verkrijgen. Het beleid is er altijd op gericht geweest om de leerlingen naar de zorg te brengen (i.c. de speciale school) in plaats van de zorg naar de leerling. WSNS wil de regelgeving zodanig aanpassen dat dergelijke wensen kunnen worden gerealiseerd. Verder, en dat is een centraal punt in de WSNS-gedachte, moet niet op één landelijk duaal systeem worden gemikt met een beperkte keuzevrijheid (handhaving in de basisschool zonder faciliteiten voor speciale zorg of plaatsing in de speciale school) maar moeten er meerdere opties worden nagestreefd (tijdelijke hulp buiten de klas, opvang in een speciale klas, extra hulp door een tweede leerkracht in de reguliere klas, intensieve cursussen voor kleine groepen, etc.). De bedoeling is dus om een verscheidenheid aan opties te ontwikkelen.

Verder, en dat is ook een centrale WSNS-gedachte, moet de keus voor een bepaalde optie zo onafhankelijk mogelijk worden gemaakt van de aanbieders van die opties. De werkers in de speciale school en andere daaraan verbonden deskundigen moeten niet worden gestimuleerd om tot verregaande vormen van klantenbinding over te gaan. Hun overwegingen moeten zoveel mogelijk staan in het belang van de creatie van een effectieve en efficiënte zorg, zo mogelijk in het basisonderwijs. Een financiering van het aanbod, i.c. de speciale school op grond van het aantal toegelaten leerlingen, is dus uit den boze. Ik heb betoogd dat ook een cliënt-gebonden budget nadelen bevat (en dit is bekrachtigd door Van Gemert) en dat het toe-

kennen van een budget op bovenschools niveau (het samenwerkingsverband) een alternatief kan zijn, mits de bestemming van die middelen is geoommerkt en er verantwoording plaatsvindt.

De beslissing over de inzet van de verschillende opties moet dus niet liggen bij de aanbieder van die zorg (noch bij een commissie die sterk met die aanbieder is geassocieerd), maar moet in samenspraak met hulpvrager (ouders, school voor basisonderwijs) geschieden, waarbij de opties breder worden opgevat dan het genoemde duale stelsel. Gestreefd wordt naar een collegiaal model, waarbij werkers in scholen voor basisonderwijs en in scholen voor speciaal onderwijs zoveel mogelijk adequate hulpvormen ontwikkelen (zo mogelijk binnen het basisonderwijs), zonder daarbij gehinderd te worden door een te strakke regelgeving waarbinnen tot nu toe in feite alleen maar de opvang in de speciale school wordt gefinancierd. En het is zeer de vraag of het merendeel van de klachten een dergelijke benadering vraagt.

**Annual Meeting American
Educational Research
Association
18-22 april 1995
San Francisco**

Inleiding (J. N. Streumer, Universiteit Twente)

De jaarlijkse conferentie van de American Educational Research Association heeft in 1995 minstens in één opzicht een record gebroken: meer dan 11.000 bezoekers werden geregistreerd. Dat impliceert dat ongeveer de helft van alle AERA-leden (in totaal 22.000) was afgereisd naar San Francisco. In totaal werden 1461 sessies georganiseerd. Iets meer dan de helft van de ingezonden papervoorstellen werd goedgekeurd en in het programma opgenomen. Dit wijkt niet af van voorafgaande jaren. Het programma bestond uit paper- en symposiumsessies, experimentele sessies, trainingsactiviteiten, poster sessies, presentaties voor nieuwe leden en graduate studenten, invited presentations, tentoonstellingen, off-site sessions en talrijke bijeenkomsten met een zakelijk en/of gezelligheidskarakter.

Het thema voor 1995 was: Partnership for a New America in a Global Community.

Ook in 1995 waren de Nederlanders goed vertegenwoordigd, zowel in de rol van conferentiebezoeker als paperpresentator. Nauwkeurige statistieken over 1995 wat betreft de nationaliteit van deelnemers en paperpresentatoren zijn nog niet bekend. De verwachting is echter dat het aantal presentatoren dit jaar hoger is uitgevallen dan vorig jaar. Om een idee te geven van de Nederlandse aanwezigheid in 1994 volgen hieronder enige gegevens, ontleend aan een artikel van Richard White in het augustus-september nummer van de Educational Researcher. Van een totaal van 4674 presentatoren in 1994 waren er 37 met een Nederlandse nationaliteit en 5 met een Belgische nationaliteit. Nederland eindigde wat betreft het aantal deelnemende presentatoren op de

vijfde plaats, voorafgegaan door de Verenigde Staten (4100 presentatoren), Canada (235), Australië (81), Engeland (77). België eindigde op een gedeelde 11e plaats.

Dit jaar zal de AERA-conferentie plaats vinden in New York van 8-12 april. Het thema zal zijn: Research for Education in a Democratic Society. De programmaprovoorzitter voor 1996, Ann Lieberman, zegt daarover: "In general, the theme suggests the importance of educational research to a variety of issues that deal with learning, teaching, and access to knowledge for increasing numbers of students in our own country and around the world. In particular it asks that we work to achieve a better understanding of the many new small learning communities being created today that are successfully engaging a diverse population of students. And further, that we look carefully at small- and large-scale efforts that are attempting to transform relationships between district and school, school and family, teacher and student, and researcher and practitioner (Lieberman, 1995 in: Educational Researcher, June-July 1995, p. 43).

De papervoorstellen voor dit jaar zijn inmiddels al weer becommentarieerd en de inzenders zijn op de hoogte van het feit of hun paper is geaccepteerd en geprogrammeerd. Zal de trend van de afgelopen jaren wat betreft een groeiend aantal Nederlandstalige presentaties en deelnemers zich ook dit jaar weer voortzetten?

Beroeps- en bedrijfsopleidingen (M. Mulder, Universiteit Twente)

Het gebied van de beroeps- en bedrijfsopleidingen wordt binnen de AERA niet afgedekt door een volwaardige divisie. Er zijn diverse Special Interest Groups (SIGs) op het gebied van de bedrijfsopleidingen, het beroepsonderwijs, opleidingen binnen defensie, opleidingen binnen de bedrijfskunde en de volwasseneneducatie. Aangezien ik als programmaprovoorzitter verantwoordelijk was voor het programma van de SIG Training in Business and Industry zal ik vooral verslag doen over de bijdragen van deze SIG. Het programma bestond uit vier sessies: twee gewone papersessies, een sessie voor zogenaamde roundtables waar geïnteresseerden intensief van gedachten kunnen wisselen over

een paper van de onderzoeker, en een business meeting waarin A. Tuijnman een inleiding zou houden over de relatie tussen opleiding en economische ontwikkeling, maar helaas moest hij enige tijd voor de conferentie afzeggen.

De belangstelling voor het programma van de SIG is al jaren vrij constant. Meestal varieert het bezoekersaantal bij de sessies van twintig tot dertig personen, met uitschieters naar veertig tot vijftig personen, wat voor een SIG redelijk goed genoemd kan worden. Ervaren bezoekers van de AERA weten dat er soms bijeenkomsten van divisies plaatsvinden in zalen met een capaciteit van honderden personen, terwijl er dan vier of vijf inleiders verslag doen van hun onderzoek voor nog geen tien bezoekers in de zaal. Ook in 1995 was dat weer het geval.

De onderwerpen van de sessies in de SIG waren 'de positie van evaluatie, team oriëntatie en klanttevredenheid', 'evaluatie van Electronic performance systems, training courseware en management training', en 'arbeidsbetrokkenheid, evaluatie en kwaliteit van opleidingsprogramma's'. Zoals zo vaak tijdens internationale wetenschappelijke conferenties, was ook binnen dit programma sprake van weinig coherentie. De onderwerpen die aan de orde kwamen, waren zeer uiteenlopend en zeer verschillend van kwaliteit.

Ik beperk me bij het bespreken van de papers tot enkele naar mijn oordeel redelijk goede. Daarbij laat ik mijn eigen bijdrage aan het programma buiten beschouwing. Ik laat het liever aan anderen over om daar een kwaliteitsoordeel over uit te spreken.

Allereerst het onderzoek van Gauthier van de Universiteit van Montreal. Zij heeft een onderzoek gedaan naar de transfer van de resultaten van een opleiding naar de werksituatie van managers en medewerkers van een instelling in de dienstensector. Dit is een belangrijk onderzoeksthema omdat vaak wordt gesteld dat maar tien tot twintig procent van de inhoud van opleidingen leidt tot werkelijke transfer op de werkplek. Het onderzoek borduurt voort op een conceptueel raamwerk van Baldwin en Ford, waarbij naast opleidingscondities, persoonskarakteristieken en kenmerken van de werkomgeving worden gezien als factoren die leren en retentie alsmede korte en lange termijn transfer beïnvloeden. Het onderzoek van Gauthier richt zich op de vraag of het organisatieklimaat en de

betrokkenheid bij het werk van invloed zijn op de transfer. In het onderzoek is gebruik gemaakt van vier experimentele groepen en een controle-groep ($n = 75$). Bij de verschillende experimentele groepen zijn op verschillende momenten gegevens verzameld. Als methoden zijn gebruikt: een observatie door externe observatoren van het werk twee maanden voor en tien maanden na de opleiding en vragenlijsten die zijn ingevuld door de deelnemers aan de opleiding. Het gaat om vier vragenlijsten voor het meten van respectievelijk de beheersing van de opleidingsdoelen, het werkgedrag in de werksituatie, de betrokkenheid bij het werk en het organisatieklimaat. Uit de resultaten van het onderzoek, die overigens niet zo helder worden beschreven, blijkt dat er sprake is van leren, retentie en opleidingseffecten, en dat betrokkenheid van het werk en organisatieklimaat inderdaad van invloed zijn op de transfer.

Vervolgens het onderzoek van Russ-Eft en Ravishankar van Zenger Miller in San Jose. Zij onderzochten de houding van managers en medewerkers ten opzichte van hun werk en vergeleken daarbij organisaties naar de mate waarin zij verschillen in hun oriëntatie op teams. In veel organisaties wordt gewerkt aan teamvorming, omdat veelal gestipuleerd wordt dat het werken in teams als resultaatverantwoordelijke eenheden leidt tot een grotere efficiëntie en effectiviteit van het werk. De onderzoekers hebben 81 managers en 364 medewerkers uit tien organisaties een oordeel laten geven van hun functie en organisatie. Daarnaast hebben onafhankelijke experts op het gebied van zelfgestuurde werkteams een oordeel gegeven over de mate waarin de organisatie is gericht op het werken in teams. Deze gerichtheid van organisaties duiden de onderzoekers aan met de term 'team oriëntatie'. Uit de resultaten van het onderzoek blijkt dat de team oriëntatie van organisaties inderdaad samenhangt met de houding van managers en medewerkers ten opzichte van de eigen functie en organisatie. Bovendien blijkt dat in teamgerichte organisaties de oordelen van managers en leidinggevendenden over hun functie en de organisatie meer met elkaar overeenkomen dan in organisaties met een lagere mate van team oriëntatie. De onderzoekers concluderen dat team oriëntatie een belangrijk middel is om het klimaat in organisaties te verbeteren. De vraag is echter of op

grond van het onderzoek wel gesproken kan worden over een causale relatie tussen team oriëntatie en organisatieklimaat. Er zijn mijns inziens evenzovele gronden om aan te nemen dat het organisatieklimaat van invloed is op de mate van team oriëntatie. Als antwoord op die vraag zou het interessant zijn om na te gaan in welke mate opleidingen in team oriëntatie van invloed zijn op het verbeteren van het organisatieklimaat.

Tenslotte het onderzoek van MacDonald, Bradley Cousins, Bailetti en Rahman van de Universiteit van Ottawa en de Carleton Universiteit. Zij hebben een model geconstrueerd voor betekenisvolle evaluatie van managementopleidingen in ondernemingen met een sterk technologisch karakter. Zij zien dit model als een alternatief voor de vele evaluaties waarbij alleen wordt ingegaan op de reacties van de deelnemers of waarbij de verhouding tussen kosten en baten wordt geïnventariseerd. Het uitgangspunt van de onderzoekers is dat eerst moet worden nagegaan wat de factoren zijn die het succes bepalen van produktontwikkelingsprojecten. Zij hebben daarvoor tien managers uit negen verschillende ondernemingen een lijst van 47 succesfactoren voorgelegd. Deze personen hebben die lijst beoordeeld met behulp van een kaartsorteertechniek en importantieweging. De gegevens zijn verwerkt met behulp van multi-dimensionale schaaltechnieken en een clusteranalyse. Het resultaat is een model met karakteristieken van individuen, groepen en werkobjecten die succesvol zijn bij produktontwikkelingsprojecten. Dit model kan worden gebruikt bij evaluaties van managementopleidingen in technologisch geörienteerde ondernemingen.

Er zullen dit jaar voor het eerst proceedings verschijnen van de papers die zijn gepresenteerd in het programma van de SIG Training in Business and Industry. Deze proceedings zullen worden uitgegeven door de Faculteit der Toegepaste Onderwijskunde onder de titel 'AERA Research Papers on Training in Business and Industry 1995'.

De SIG op het gebied van het beroepsonderwijs is voornemens een divisie op te richten getiteld 'Education and Work'. Initiatiefnemer is prof. dr. Curtis Finch. Het is de bedoeling om niet te

concurreren met de bestaande SIGs, ofschoon de oprichting van deze divisie wel de nodige effecten zal hebben op de SIGs. Veel leden van verschillende SIGs zullen namelijk belangstelling hebben voor deze nieuwe divisie. Hoe dan ook, de intentie is de divisie op te richten naast de bestaande SIGs, zodat deze hun eigen activiteiten gewoon kunnen voortzetten.

Vergeleken met de Nederlandse situatie binnen de Vereniging voor Onderwijsresearch zou deze AERA-divisie verwant zijn aan de VOR-divisie Beroeps- en Bedrijfsopleidingen. Ik ondersteun het initiatief om te komen tot deze divisie omdat ik denk dat het veel voordelen biedt voor onderzoekers op het terrein van de beroeps- en bedrijfsopleidingen. Nu is hun onderzoek sterk versplinterd in allerlei SIGs binnen de AERA. De beoogde divisie kan waarschijnlijk meer synergie in de presentatie van onderzoeksresultaten bewerkstelligen. Voorwaarde voor het oprichten van een divisie binnen de AERA is echter dat een substantieel deel van de AERA-leden het initiatief ondersteunt. Geïnteresseerde VOR-leden van de divisie Beroeps- en Bedrijfsopleidingen kunnen zich voor verdere informatie melden bij de auteur van dit deel van de AERA-Kroniek. Hij treedt op als contactpersoon voor de Amerikaanse initiatiefnemer.

Instructie- en Communicatietechnologieën (J.M. Pieters, Universiteit Twente)

De tijd dat er op de AERA presentaties werden verzorgd waarbij de onderzoeker zijn of haar resultaten presenteerde met handgeschreven en/of moeilijk leesbare transparanten lijkt voorgoed voorbij. De presentaties zijn goed verzorgd en de moderne media hebben hun intrede gedaan. Het gebruik van Power Point of vergelijkbare presentatiesystemen, off-line dan wel on-line, is vanzelfsprekend. Bij een aantal sessies zag ik de discussiant tijdens de presentaties zijn of haar discussie voorbereiden op de computer, vanzelfsprekend van notebook formaat. De discussie zelf was daardoor zeer to-the-point en visueel van hoog niveau. Ook de dynamica heeft zijn intrede gedaan. Tijdens de presentatie dynamische voorstellingen het betoog laten ondersteunen, soms zelfs met aansluiting op een externe computer, komt ook voor.

Is er dan ook aan de inhoud van de presentaties ook een en ander veranderd? Jazeker, de nieuwe technologieën bieden mogelijkheden op ruimere schaal ingenieuzer te experimenteren, gegevens te verzamelen en te verwerken, maar vooral de leeromgeving van een nieuwe dimensie te voorzien, namelijk interactiviteit. Interactiviteit kan zich op verschillende manieren openbaren: in kennisnetwerken, in kennisconstructie met een scala aan participanten, in collaboratie, en in situationisme en realisme.

Wanneer we de presentaties in Divisie C overzien dan constateren we een overwegend gebruik van termen als: constructivism, interacting, (learning) community, knowledge organization, communication, computer-mediated instruction, learning environments, discourse communities, en virtual communities. Een goed voorbeeld van het gebruik van nieuwe technologieën in het onderwijs (in constructivistische zin) gaf het symposium met de titel 'Applications in networked multimedia for knowledge construction, collaboration, and curriculum reform'. Een sterke nadruk lag in dat symposium op de technische toepassingsmogelijkheden waarbij de theoretische en in het verlengde daarvan empirische onderbouwing het moest afleggen tegen de aantrekkingskracht van het nieuwe medium, ondanks het gebruik van de termen knowledge construction en collaboration. Ook in het symposium 'The Learning Community: A Star schools project building a network of school-community partnership through multiple technologies' waarin verslag gedaan werd van de eerste ervaringen in dit grootschalige project, overheerste het imponerende karakter van de nieuwe technologieën en optimistische visies op een succesvolle implementatie. Deze door de technologie gedomineerde aanpak, eerst de technische implementatie en pas daarna je druk maken over de theoretische verantwoording, stak schril af tegen de wat meer pessimistische, misschien ook wel meer realistische kijk op de zaak van door de onderwijskundig-theoretische en -praktische wol geverfde onderzoekers als Fullan, Rothkopf en Berliner die in het symposium 'Getting instructional research results in our schools: A loveless marriage, temporary misunderstanding, or promising partnership' de traditionele R&D benadering (met eigentijdse variant) verdedigden. Een tussenpositie

nemen de onderzoekers in die in het symposium 'Schools for thought: Transforming classrooms into learning communities' vanuit zeer geprononceerde theoretische opvattingen fraaie praktische telematische innovaties presenteerden. Innovaties die naar de mening van de onderzoekers, gesteund door weliswaar beperkte empirische evaluaties, er wel degelijk toe deden. Dat konden deze onderzoekers elkaar en het publiek in een aantal symposia vol overtuiging duidelijk maken. Het gaf ten langen leste de indruk van een rondreizend circus. De onderzoekers, in paren: Brown en Campione met hun Communities of practice, Scardamalia en Bereiter met CSILE, en Goldman en Bransford met Jasper. Het moet gezegd: de innovaties zijn theoretisch en praktisch indrukwekkend, maar waarom moet dat in een flink aantal symposia telkenmale herhaald worden zonder dat er veel nieuwe feiten aan worden toegevoegd?

Een interessante ontwikkeling die ook op de AERA zichtbaar werd in de paper-presentaties, is de verbinding van communicatie en interactie. De nieuwe technologieën, in het bijzonder een nieuwe communicatietechnologie als Internet, maakt het mogelijk geheel nieuwe en tot voor kort nauwelijks denkbare innovaties op het terrein van instructie en leren door te voeren. De lerende wordt steeds minder afhankelijk van de produkten van de instructieontwerper maar wordt meer in staat gesteld in rijke omgevingen, met behulp van de communicatietechnologie, rond te dolen en daardoor zijn eigen kennispakket samen te stellen. Symposia en presentaties met in de titel knowledge building of knowledge construction, met of zonder collaborative, maakten hier steeds meer melding van. Zoals bijvoorbeeld het door OISE gedomineerde symposium 'Collaborative knowledge building on a computer network', en de interessante paper sessie 'Technology enhanced learning environments' met bijdragen van Woodruff 'The effects of computer-mediated communication on collaborative discourse in knowledge building communities' en van Jonassen over case-based learning environments. Ook de bijdragen van Levin (Illinois) in een aantal symposia, waarin hij soms optrad als discussiant, wezen duidelijk op mogelijkheden in die richting.

Maar ondanks de indruk die presentaties en

demonstraties maken, er blijft de vraag naar de effectstudies (wat is nu het effect van die mooie omgevingen?) en naar de ontwerprichtlijnen voor die fraai vormgegeven omgevingen. Sommige onderzoekers trachten daarop antwoorden te vinden door communicatienetwerken te observeren en de interacties tussen docenten, domein-experts en studenten systematisch te inventariseren en zodoende een beeld te schetsen van de werking van niet alleen de technologische omgeving, maar met name van de inhoudelijke impact ervan. Een voorbeeld van het laatste werd geleverd in het symposium 'Analyses of computer-mediated communications within on-line teaching/learning' door Harris en Jones in hun bijdrage 'A study of on-line communication among subject matter experts, teachers, and students: Message flow and functions'.

Professionele ontwikkeling van leraren (M. G. Roes, Universiteit Twente)

Het realiseren van onderwijsvernieuwingen is en blijft een complexe aangelegenheid. Dit bleek weer eens uit de papersessies waarin verslag werd gedaan van Amerikaans onderzoek naar de effectiviteit van programma's voor de professionele ontwikkeling van leraren.

Leraren in de VS zien zich voor ingrijpende onderwijsvernieuwingen geplaatst: nieuwe benaderingen van de exacte vakken, nieuwe instructiemodellen als 'complex instruction' en coöperatief leren. Deze innovaties hebben met elkaar gemeen dat ze uitgaan van een constructivistische opvatting van het leerproces.

De ondersteuning van docenten bij de invoering van deze vernieuwingen wordt intensief aangepakt. Langer lopende trajecten bestaande uit cursussen, follow-up workshops en begeleiding op de werkplek in de vorm van coaching, zijn geen uitzondering.

In veel van de gepresenteerde studies waren de 'knowledge and belief systems' van leraren als uitgangspunt genomen voor het inrichten van de professionele ondersteuningsarrangementen. Via veranderingen in kennis en opvattingen van docenten worden veranderingen in de lespraktijk nagestreefd. In de uitwerking van deze programma's betekent dit dat veel belang wordt gehecht aan mogelijkheden voor continue dialoog met collega's en programma-

staf én voor reflectie op de eigen lespraktijk.

Een voorbeeld hiervan was het paper van Swafford, Jones en Thornton met als titel 'Changing middle grades teachers' classroom practices'. Dit paper werd gepresenteerd tijdens een sessie waarin de vernieuwing van het wiskunde-onderwijs centraal stond. De onderzoekers hadden de effecten bestudeerd van een driejarig ondersteuningsprogramma dat gericht was op de implementatie van nieuwe standaarden voor het wiskunde-onderwijs. Het nascholingsprogramma bestond jaarlijks uit de volgende componenten: een zomercursus wiskunde van vier weken; een ééndaags research-seminar over het leren van wiskunde door leerlingen én zes middagseminars, waarbij gewenste lespraktijken werden geanalyseerd, onder meer op basis van inbreng van de deelnemers. Ook was tijdens deze middagen gelegenheid voor formele en informele uitwisseling tussen de deelnemers. Als huiswerk moest iedere deelnemer jaarlijks twee lessen op video opnemen en analyseren en een reflectie-logboek bijhouden. Aan het eind van elk 'cursus'jaar moesten ze een samenvatting van deze logboeken maken en bespreken met de programmaleiding. Tenslotte was er op de scholen van de deelnemers tijd voor onderlinge samenwerking uitgetrokken. Om de effecten van dit programma te onderzoeken, hadden de onderzoekers de 48 deelnemers aan het begin en aan het eind van het traject een vragenlijst voorgelegd over hun opvattingen, lespraktijk en leerstofaanbod en een wiskundetoets afgenomen. Zes deelnemers waren twee keer per jaar een week lang geobserveerd en geïnterviewd. Bij de leerlingen van alle deelnemers (en bij leerlingen in een controlegroep) was een wiskundetoets en een attitudevragenlijst afgenomen. De onderzoekers constateerden op alle onderzochte aspecten significante veranderingen in de gewenste richting. Uitzondering hierop vormde de attitude van de leerlingen van de deelnemers, die niet verschilden van de leerlingen in de controlegroep. De onderzoekers concluderen dat vernieuwingen in het wiskunde-onderwijs mogelijk zijn, maar dat het veel tijd kost: pas na het eerste jaar werden veranderingen zichtbaar.

Ondanks de hoopgevende resultaten van deze studie, bleek het effect van de diverse programma's voor de professionele ontwikkeling

van leraren nogal wisselend te zijn. Bestaande opvattingen van leraren over onderwijzen blijken vaak zeer hardnekkig te zijn. Deze opvattingen passen vaak meer bij een transmissie-model van het onderwijsleerproces dan bij een constructivistische benadering.

In de papers lag overwegend de nadruk op de kennis en opvattingen van docenten als aangrijpingspunt voor veranderingen. Het in de praktijk brengen van nieuwe modellen van onderwijzen, vraagt echter ook om de verwerving van nieuwe vaardigheden. Op de conferentie was nauwelijks aandacht voor gerichte vaardigheidstraining als component van een programma voor professionele ondersteuning. Evenmin werd aandacht geschonken aan de rol die curriculummaterialen kunnen spelen in een dergelijk traject. Wellicht dat een meer integrale benadering de effectiviteit en efficiency van dergelijke ondersteuningsprogramma's kan vergroten.

Schooleffectiviteitsonderzoek (H. de Vos, Universiteit Twente)

Alhoewel er op de AERA in 1995 zo'n 1400 papersessies waren, werden er maar drie gekenmerkt als schooleffectiviteit-sessies. Eén ging over de relatie van schooleffectiviteitsonderzoek en schoolverbetering, één over onderwijs indicatoren en de laatste ging over voortgang van schooleffectiviteitsonderzoek. De sessies kenmerkten zich door hun internationale karakter: de presentaties werden niet alle door Amerikanen verzorgd.

Aangezien mijn interesse voornamelijk methodologisch was getint, zal ik met name de presentaties eruit lichten, die aan mijn interesse tegemoet komen.

Een geheel Nederlandse sessie was verroosterd op donderdagmorgen: The evaluation of the Dutch Educational Priority Program. Het publiek bestond grotendeels uit Nederlanders. Het onderwerp was, voor zover dat nog niet duidelijk was uit de titel, de evaluatie van zes jaren OVB. De participanten aan deze sessie waren Van Tilborg en Walraven van SVO, Mulder van het ITS en Van der Werf van het GION. Alhoewel blijkt dat de prestaties op de OVB-scholen verbeterd zijn, is het moeilijk te zeggen of dit ook door het OVB is bewerk-

stelligd. De oorzaak zou ook kunnen liggen in het feit dat de leerlingen zes jaar langer in Nederland wonen. Welk deel van de verbetering wordt verklaard door factoren binnen de school, factoren buiten de school en OVB is nog onduidelijk. De discussiant Slavin, van de John Hopkins University, was erg enthousiast over de OVB, mede omdat het Nederlandse onderwijsbeleid ten aanzien van specifieke doelgroepen soms lijnrecht tegenover de uitkomsten van het gedecentraliseerde onderwijs-financieringsstelsel van de Amerikaanse staten staat: (relatief) arme staten besteden minder geld aan de compensatie van achterstanden van gedepriveerde groepen leerlingen.

Ook donderdagochtend (om 8.15 uur!) was er een sessie over internationaal gebruik van onderwijs-indicatoren: International uses of quality indicators. De toevoeging "internationaal" was zeker op zijn plaats; bijdragen aan deze sessie kwamen uit Canada, Engeland, Frankrijk, Nederland en de Verenigde Staten. Onderwijsindicatoren worden gebruikt om de kwaliteit van onderwijs en scholen te beschrijven. De papers belichtten zienswijzen omtrent ontwikkeling en optimaal gebruik van deze indicatoren. In Canada, aldus McEwen, bepalen de provincies het onderwijs en dus ook het gebruik van indicatoren. Zij stelde een aantal verbeteringen voor op provinciaal en nationaal niveau, zoals het gebruik van indicatoren voor vergelijkingen tussen provincies (een soort benchmarking?) en voor het optimaliseren van schoolbegeleiding. Uit Engeland kwam een bijdrage van Fitz-Gibbon. Zij vroeg zich af of het systeem van indicatoren dat wordt gebruikt in de UK, wel een beeld geeft van de kwaliteit van onderwijs. Het indicatorensysteem wordt met name bepaald door examenresultaten en resultaten van onderwijsinspectie. De kwaliteit van het indicatorensysteem wordt momenteel verhoogd door de toevoeging van 'value-added' maten. Creemers (Nederland) ging in op de bijdrage van schooleffectiviteitsonderzoek aan de ontwikkeling en het gebruik van onderwijsindicatoren. Hij stelde dat onderzoek, zoals IEA (International Association for the Evaluation of Educational Achievement) en ISERP (The International School Effectiveness Research Program), de ontwikkeling ondersteunen. Het is echter wel zaak, aldus

Creemers, dat indicatoren internationaal vergelijkbaar zijn, dus onafhankelijk van het onderwijsstelsel. Bonnet uit Frankrijk hield een pleidooi voor het monitoren van het onderwijs met behulp van indicatoren. Hij waarschuwde dat moet worden opgepast dat indicatoren geen politieke instrumenten worden.

Op de vrijdagmiddag was er een sessie over internationale voortgang in schooleffectiviteitsonderzoek georganiseerd door ISERP, onder voorzitterschap van Reynolds. Tymms (UK) ging in zijn presentatie 'Monitoring effectiveness: Examples of the evaluation of complex systems' ver door te beweren dat schooleffectiviteitsonderzoek er nooit in zal slagen zodanige kennis te vergaren dat scholen in het algemeen kunnen worden verbeterd, omdat wat er in scholen en klassen gebeurt onvoorspelbaar, complex en chaotisch is. Hij hield een pleidooi voor gebruik van onderwijsindicatoren, die meer inzicht kunnen geven in wat er nu eigenlijk gebeurt. In de paperpresentatie over 'Two year quantitative findings from ISERP' (o.a. Creemers) werd ingegaan op het gegeven dat de leerlingvariabelen (SES, IQ, ethniciteit) het grootste deel van de wiskundeprestaties van leerlingen verklaren in internationaal vergelijkend onderzoek (net als in het meeste onderzoek). Alhoewel in de meeste landen ook scholenfactoren van invloed zijn op de leerlingprestaties is dat in Nederland (en ook in Taiwan) niet het geval.

Measurement and Research Methodology (H. J. Vos, Universiteit Twente)

Het aandachtsgebied 'Measurement and Research Methodology' was ruim vertegenwoordigd in de vorm van thematische papersessies, (invited) symposia waarvan 36 in samenwerking met de National Council on Measurement in Education (NCME), posterpresentaties en pre-conference workshops.

Een van de belangrijkste trends betrof het adaptief toetsen van vaardigheden. Computerized Adaptive Testing (CAT) wordt steeds meer gebruikt in de praktijk en lijkt de toetsvorm van de toekomst te worden. Het belangrijkste voordeel van deze toetsvorm in vergelijking met conventionele potlood-en-papier (P&P) toetsen is dat volstaan kan worden met kortere toet-

sen bij gelijkblijvende of zelfs hogere nauwkeurigheid van de geschatte vaardigheid. Er is echter nog veel psychometrisch onderzoek te doen op dit gebied, zowel voor 'fine-tuning' van de theorie als voor het oplossen van praktische problemen bij de implementatie. Aan beide aspecten van CAT werd in verschillende papersessies en symposia ruimschoots aandacht besteed.

Veerkamp en Chang stelden beiden nieuwe criteria voor om items te selecteren in een adaptieve toetsprocedure gebruikmakend van item response theorie (IRT) modellen. Het belangrijkste bezwaar bij de meeste huidige criteria is dat ze geen rekening houden met de onzekerheid van de geschatte vaardigheid. Terwijl Veerkamp een gewogen gemiddelde van de gebruikelijke informatiefunctie voorstelde, stelde Chang daarentegen een 'global information approach' voor om de gesignaleerde bezwaren bij de huidige item selectiecriteria te omzeilen.

In de presentatie van Rocklin, Vispoel en Wang over de mogelijkheid tot manipulatie van het CAT-algoritme als item review mogelijk is door de 'test-taker' (Self-Adapted Testing) kwam naar voren dat het erg riskant is voor de test-taker om op deze manier te trachten onterecht een hogere score te behalen. In theorie kan een test-taker door er voor te zorgen alleen erg gemakkelijke items voorgelegd te krijgen en die vervolgens allemaal goed te beantwoorden kunstmatig een erg hoge score verkrijgen. Als echter enkele vragen fout beantwoord worden zal zijn of haar score te laag uitvallen.

Lunz en DeVille deden verslag van een studie waarin toetsen geconstrueerd via de computer met een adaptieve selectieprocedure werden vergeleken met toetsen die handmatig werden geconstrueerd en op papier werden afgenomen. Hierbij werd met name gekeken naar de validiteit van de geselecteerde items, de volgorde van de items, de statistische eigenschappen van de toetsen en in hoeverre de geselecteerde items voldeden aan de toetsspecificaties. Uit deze studie bleek dat beide methoden voor toetsconstructie vergelijkbaar zijn voor wat betreft validiteit en statistische eigenschappen en dat beide methodes voldeden aan de toetsspecificaties.

Davey en Parshall keken naar exposure

rates van items (percentage test-takers dat een bepaald item voorgelegd krijgt) in adaptieve toetsen. Er werd gekeken hoe hoog deze exposure rates zijn als er niet voor gecontroleerd wordt. Verschillende methodes van item exposure control werden voorgesteld en met elkaar vergeleken.

Schnipke onderzocht hoe itemresponstijden gebruikt kunnen worden om te achterhalen hoe een persoon een item heeft trachten te beantwoorden, door snel te gokken of door serieus naar een oplossing te zoeken. Op deze manier is het mogelijk gokgedrag uit de test te filteren en daardoor de vaardigheid beter te schatten.

Tenslotte presenteerden Way, Smith en Lewis hun paper over het gebruik van Bayesiaanse besliskunde bij 'computerized mastery testing' (CMT). Testlets (d.w.z. een klein aantal samenhangende items) werden sequentieel afgenomen in hun onderzoek. Na iedere aangeboden testlet werd een slaag, zak, of continueer beslissing genomen. Van duidelijke masters of nonmasters kon na slechts enkele testlets reeds worden besloten of ze waren geslaagd of gezakt. Test-takers wier prestaties echter in de buurt van de grensscore lagen moesten de meeste testlets beantwoorden.

Zoals gebruikelijk de afgelopen jaren gingen veel papers over IRT-modellen en statistische problemen hierbij die nog om een oplossing vragen. Het symposium 'Statistical issues in item response theory models', georganiseerd door Berger, was hiervan een goed voorbeeld.

Verhelst rapporteerde over het door het CITO ontwikkelde gegeneraliseerde One-Parameter Logistic Model (OPLM). Hij liet zien dat bij het OPLM-model enerzijds de elegante statistische eigenschappen van het Rasch-model bewaard blijven terwijl anderzijds de grotere flexibiliteit van het 2-parameter logistische model (Birnbaum model) optimaal kan worden benut.

Glas toonde aan dat in een klasse van IRT-modellen, welke het Rasch model, het lineair logistische test model (LLTM), het (gegeneraliseerde) partial credit model, het OPLM-model en modellen met multidimensionale persoonsparameters bevatte, zowel Conditional Maximum Likelihood (CML) als Marginal Maximum Likelihood (MML) schattingsprocedures kunnen worden toegepast. Tevens werd een taxonomie van gegeneraliseerde

Pearson tests beschreven welke power hebben tegen specifieke modelschendingen in deze klasse van modellen. De benadering werd geïllustreerd aan de hand van twee voorbeelden om item bias (DIF) te detecteren.

Sijsma deed verslag over nonparametrische IRT-modellen voor dichotome en polytome data, welke een ordinale ordening van personen en/of items veronderstellen. Hij vergeleek nonparametrische en parametrische IRT-modellen met elkaar en ging hierbij met name in op de vraag in welke situaties het gebruik van nonparametrische IRT-modellen de voorkeur genieten (o.a. wanneer de assumpties bij logistisch parametrische modellen te restrictief zijn).

Van der Linden en Luecht presenteerden een optimalisatiemodel (0-1 lineair programmeringsmodel) dat toetsconstructeurs in staat stelt om de vorm van een vooraf gekozen scoreverdeling op een toets voor een populatie met een bekende vaardigheidsverdeling zo goed mogelijk te controleren.

Tenslotte toonde Kelderman in een meer fundamenteel paper aan dat het Rasch model kan worden afgeleid uit een gegeneraliseerd lineair model voor de uitwisselbaarheid van metingen.

Een laatste opvallende trend was de hernieuwde interesse voor het probleem van het vaststellen van grensscores. Na een hausse in de jaren zeventig en tachtig als gevolg van mastery learning was tijdens de afgelopen AERA/NCME dit thema weer prominent aanwezig op de onderzoeksagenda. Het symposium 'Applications of decision theory to practical testing problems', georganiseerd door Sawyer met Van der Linden als moderator en vijf bijdragen omvattend, ging met name in op de rol die Bayesiaanse besliskunde kan spelen bij het vaststellen van grensscores.

Ben-Shakhar, Kiderman en Beller deden verslag van een vergelijkende studie naar twee verschillende beslissingstheoretische toelatingsprocedures voor 'Liberal Arts' op universitair niveau. De ene toelatingsprocedure hield wel rekening met eerder behaalde SAT (Scholastic Aptitude Test)-scores, terwijl de andere procedure hier geen rekening mee hield. De resultaten toonden aan dat er geen winst te behalen valt bij toevoeging van de SAT-scores als het de bedoeling is om studenten te recruter

die voldoen aan minimale kwalificaties. Wanneer het echter de bedoeling is om potentieel excellente studenten te selecteren bleek toevoeging van de additionele informatie van de SAT-scores aanzienlijke winst op te leveren.

Van der Linden presenteerde een beslissingstheoretisch model voor het veel in de praktijk voorkomende probleem van het optimaal toewijzen van studenten aan twee verschillende hiërarchisch geordende cursussen ('course placement'). Op basis van een aptitude test moest worden besloten of studenten ofwel werden toegewezen aan de eerste cursus ofwel rechtstreeks werden toegelaten tot de tweede cursus. De eerste cursus werd hierbij afgesloten met een beheersingstoets welke bij een voldoende resultaat toegang gaf tot de tweede cursus.

Vos rapporteerde over het simultaan optimaliseren van combinaties van onderwijskundige beslissingen. De belangrijkste toepassingen hiervan liggen op het terrein van het formuleren van optimale regels voor voortgangsbepalingen in geïndividualiseerde instructiesystemen. In een empirisch geïndividualiseerd instructiesysteem, bestaande uit een gecombineerde selectie-plaatsings-beheersingsbeslissing, bleek dat bij de simultane benadering een aanzienlijke winst in verwachte utiliteit te behalen viel vergeleken bij het afzonderlijk optimaliseren van iedere beslissereg. Tevens werd aangetoond dat onder tamelijk milde voorwaarden optimale beslisseregels noodzakelijkerwijs compensatorisch van aard waren.

Terwijl de drie bovenstaande bijdragen alle uitgingen van bekend veronderstelde utiliteitsfuncties hielden de laatste twee bijdragen zich bezig met het empirisch specificeren van utiliteitsfuncties. Als zodanig vormden ze een goede aanvulling op de eerdere papers, omdat het voor een beslissingstheoretische benadering essentieel is te kunnen beschikken over een realistische utiliteitsstructuur.

Mellenbergh, Van der Gaag en Van den Brink lieten zien dat empirische utiliteitsfuncties voor cultuureerlijke selectie bevredigend konden worden benaderd door lineaire- en normaal-ogief utiliteitsfuncties.

Tenslotte toonde Sawyer aan dat het empirisch specificeren van utiliteitsfuncties voor 'course placement' tot aanzienlijke verschillen

in uitkomsten kan leiden afhankelijk van het feit of een deterministische of stochastische (Von Neumann-Morgenstern) techniek werd gebruikt.

Aan deze kroniek werkten mee: *M. Mulder, J. M. Pieters, M. G. Roes, J. N. Streumer, H. J. Vos en H. de Vos.*

De eindredactie werd verzorgd door *J. N. Streumer.*

Boekbesprekingen

J. Lowyck & N. Verloop (Red.)

Onderwijskunde: een kennisbasis voor professionals

Wolters-Noordhoff, Groningen 1995.

352 pagina's f 74,-, ISBN 90 0154 531 9

De schrijvers stelden zich minstens twee doelen met het schrijven van het boek. Allereerst een opvolger te maken voor het bekende standaardwerk 'Beknopte Didaxologie'. Vervolgens dient het boek opnieuw een "baken te zijn in de steeds groeiende stroom van sociaal-wetenschappelijke literatuur betreffende onderwijzen en leren" (citaat uit De Corte et al., 1981, p. V).

Het boek is voor de volgende gevarieerde doelgroep geschreven: a) studenten onderwijs-wetenschappen, b) afgestudeerden werkzaam in allerlei maatschappelijke sectoren die hun kennis willen opfrissen en c) leraren in opleiding of nascholing, die behoefte hebben aan een theoretische basis voor hun huidige of toekomstige werk.

Hoewel 'Onderwijskunde' als opvolger van 'Beknopte Didaxologie' wordt gepresenteerd, kunnen we het boek als een volledig nieuw produkt beschouwen. Het lijkt evenveel op zijn voorganger als een geadopteerd kind op zijn adoptief ouders. Didaxologie als begrip heeft allereerst afgedaan. Het bijbehorende ordeningskader is losgelaten. 'Beknopte Didaxologie' was opgebouwd volgens een model voor didactisch handelen, met daarin de volgende componenten: (I) doelstelling, beginsituatie, (II) het onderwijsleerproces bestaande uit de componenten didactische werkvormen, media, leerinhouden, groeperingsvormen, en (III) evaluatie en twee feedbackloops lopend van III naar I en II. Didactisch handelen als intentioneel gebeuren lijkt in dit boek wat naar de achtergrond te zijn verschoven.

Het boek is volgens de auteurs opgebouwd vanuit het perspectief van de docent. De auteurs hanteren hierbij een ruime opvatting. Men zou verwachten dat de leraar zich het meest zal herkennen in de onderdelen die direct betrekking hebben op de classesituatie en minder in de aspecten op meso- en macroniveau. In een aantal hoofdstukken worden deze laatste aspecten wel degelijk uitgewerkt. We denken daarbij aan ont-

wikkelingen in de maatschappij (Hst. 3), vernieuwingsgedachten (Hst. 3), beroepsbeelden (Hst. 4), opvattingen en basisoriëntaties over het curriculum (Hst. 6) en aspecten van kwaliteitszorg op schoolniveau (Hst. 9). Men zou misschien beter kunnen stellen dat deze inhouden gekozen zijn omdat de leraar, volgens een modern perspectief op het beroep (Commissie Van Es) naast de directe classesituatie op de hoogte zou moeten zijn van deze aspecten.

Waar het de directe rol van de leraar in de klas aangaat staat in 'Onderwijskunde' het beïnvloeden van leerprocessen meer centraal. Daarbij wordt aandacht besteed aan het type leerprocessen dat op gang gebracht wordt, het soort onderwijsstrategieën dat kan worden toegepast, wie daarbij de leerfuncties stuurt, en welke resultaten dit oplevert. Er wordt een groot belang gehecht aan het kennisconstructieproces. De taak van de leraar wordt meer en meer het creëren van rijke leeromgeving en Lowyck drukt het in hoofdstuk 7 als volgt uit: "een essentieel voorwerp van beslissingen is dan, hoe een leerling geholpen kan worden om door verschillende maatregelen in de leeromgeving juist die ondersteuning te krijgen waar hij of zij op dat moment behoefte aan heeft" (p. 242).

We kunnen over het algemeen stellen dat leerpsychologische kennis in dit boek veel meer geïntegreerd is dan bij de voorganger. In 'Beknopte Didaxologie' kwam nog een apart hoofdstuk voor over "leerpsychologie in didactologisch perspectief." In het algemeen geldt voor het boek dat de verschillende deeldisciplines meer zijn geïntegreerd in de verschillende onderwerpen. Bespiegelingen over het wat en waarom van onderwijs en opvoeden zijn ingebed in concrete ontwikkelingen in de maatschappij, het onderwijsbeleid, en de gevolgen daarvan voor de inrichting van het onderwijs (hoofdstuk 3). Het boek toont aan dat onderwijskunde als interdisciplinaire richting volwassen geworden is.

Begonnen wordt met de bespreking van een aantal 'gegevenheden': leerlingkenmerken, de school en de klas, de maatschappij en onderwijsbeleid. Vervolgens wordt de leraar zelf als factor beschouwd. Het derde deel betreft een aantal aspecten van het onderwijsleerproces die de docent kan beïnvloeden: het leerproces, realisering

van het curriculum, didactische werkvormen en media en de didactische evaluatie. Elk hoofdstuk start met een grafisch ordeningskader, waarin de belangrijkste begrippen en hun onderlinge verbanden worden weergegeven. Over het algemeen zijn de schema's verhelderend, hoewel de schema's nu eens van het type oorzaak-gevolg zijn, dan weer hiërarchisch en dan weer temporeel. Bij het schema van hoofdstuk 8, ontgaat ons de betekenis der pijlen.

In hoofdstuk 1 beschrijft Simons hoe leerprocessen en leerprestaties worden beïnvloed door verschillende leerlingenkenmerken. Veel aandacht krijgen de zogenaamde directe leerlingenkenmerken, zoals leer- en denkactiviteiten, voorkennis, de beleving van de leersituatie en de leerpsychologische kenmerken die daarop invloed uitoefenen. Als voorbeeld beschrijft Simons het onderwerp leerstijlen meer uitgebreid. Het hart van dit hoofdstuk wordt gevormd door de theorie over leervermogen (als leerpsychologisch kenmerk), waarin beschreven wordt dat bij verschillende soorten leeromgevingen een ander beroep wordt gedaan op het sturen van leerfuncties (voorbereiden, verwerken, en reguleren). Een uiteindelijk doel van onderwijs kan zijn leerlingen uiteindelijk hun eigen leerfuncties in de hand te laten nemen. Leeftijd, geslacht, thuismilieu en intelligentie worden beschouwd als variabelen die, via de leerpsychologische en directe kenmerken, hun invloed op het leren laten gelden.

Hoofdstuk 2 vormt niet echt een eenheid, omdat de onderwerpen school en klas als afzonderlijke grootheden worden behandeld. Waarom is niet gekozen voor twee afzonderlijk hoofdstukken? Marx en De Vries betogen over de trend van toenemend beleidsvoerend vermogen op scholen, die gezien kan worden als algemeen antwoord op de vele veranderingen in en rondom het onderwijs, waar scholen via beleidswijzigingen op in dienen te spelen. De auteurs gaan ervan uit dat scholen met een traditionele segmentale organisatie onvoldoende in staat zullen zijn de huidige en toekomstige ontwikkelingen op te vangen.

Omdat voor het deel over de klas weinig plaats is ingeruimd (Veenman & Slegers) heeft het een nogal opsommend karakter. De klas als leeromgeving wordt vanuit verschillende perspectieven beschreven: de fysieke leeromgeving, het klaslimaat, de groeps grootte, de klasseor-

ganisatie (i.c. groepeeringsvormen), en de dominante taakstructuren. Hierbij maakt men in ruime mate gebruik van beschikbare empirische gegevens. Het is interessant dat bij klaslimaat nieuwe ontwikkelingen in de kijk op leren worden meegenomen: in de oude opvatting stond de afstemming van individuele behoeften op maatschappelijke verwachtingen centraal. In de nieuwe opvatting wordt benadrukt dat leerlingen door interactie eigen kennis en normen construeren onder begeleiding van een leraar.

Maatschappij en onderwijsbeleid vormen het onderwerp van hoofdstuk 3 (Knoers). Dit hoofdstuk is zeer samenhangend geschreven. Gestart wordt met de beschrijving van een aantal maatschappelijke ontwikkelingen, welke uitmonden in een beschrijving van een spanningsverhouding tussen maatschappij en onderwijs. In relatie hiermee worden (veranderingen) in onderwijsopvattingen en parallel lopende overheidsmaatregelen beschreven. Ook worden achtergronden en effecten van specifieke maatregelen als deregulering, autonomisering en schaalvergroting beschreven. Er volgt een intermezzo over kwaliteit van het onderwijs, een overzicht van ondersteunende instellingen en hun invloed op de kwaliteit van het onderwijs. Voor concrete vernieuwingen als de basisschool, basisvorming en voortgezet beroepsonderwijs en zeer specifieke vernieuwingen als zorgverbreding is tevens plaats ingeruimd. In zijn betoog neemt Knoers op verschillende plaatsen nadrukkelijk stelling. Zo wijst hij op het gevaar van volledig vraaggestuurde onderwijsverzorging als er aan de vraagzijde geen sprake is van goed georganiseerde onderwijsprofessionals die onderwijskundig en pedagogisch zijn geïnteresseerd.

In het zeer degelijk geschreven hoofdstuk 4 staat de professionaliteit van het leraarschap centraal. Als achtergrond over het denken van professioneel leraarschap schetst Verloop een viertal beroepsbeelden die alle een aantal beperktheden kennen: kennisgerichte, persoonsgerichte, vaardigheidsgerichte en praktijkgerichte beroepsbeelden. Wellicht doet hij het vaardigheidsgerichte beeld wat tekort door al te nadrukkelijk te wijzen op de gevaren van fragmentarisering. In Nederland zijn de laatste 10 jaar veel succesvolle begeleidingsinitiatieven geweest waarin vaardigheden (klassenmanagement, instructie) centraal stonden. Het is in dit verband opvallend dat het werk van Veenman c.s. nau-

welijks genoemd wordt.

Tegenover eenzijdigheid binnen een bepaald beroepsbeeld stelt Verloop het alternatief van een kijk op het docentenberoep waarin nadruk gelegd wordt op een voortschrijdende ontwikkeling. Dit houdt in dat docenten lerende professionals zijn waarbij in elk stadium van hun loopbaan specifieke knelpunten en groei behoeften aan de orde zijn. Deze beginnen al met het ingroeien in het beroep, waarbij al leraren-opleiding op professionele ontwikkelscholen terecht kunnen komen. Professionaliteit blijft niet beperkt tot het individu. De kwaliteitszorg voor leraren en schoolontwikkeling zullen meer en meer op elkaar betrokken worden.

Voorwaarden voor bewaking van professionaliteit zijn dat aandacht wordt besteed aan cognities van docenten die bepalend zijn voor hun handelen, dat docenten een kritisch-reflectieve houding verwerven en dat ze een gedegen kennisbasis opbouwen. Naast vakdidactische, empirische (effectieve instructie) kennis zou dit ook praktijkkennis moeten betreffen. Misschien wordt in dit hoofdstuk het gevaar van fragmentariseren enigszins overschat gezien het succes in de begeleidingspraktijk, bijv. bij het landelijk schoolverbeteringsproject en het leren omgaan met combinatieklassen bij Veenman.

Hoofdstuk 5 (Verschaffel) is ons inziens naast 4 en 7 een van de hoogtepunten van het boek. Gestart wordt met een korte beschrijving van de ontwikkeling in theorieën over leren. Vervolgens worden voor drie leerdomeinen uitwerkingen gegeven: het cognitieve, sociaal-affectieve en psychomotorische leren. Het eerstgenoemde domein krijgt de meeste aandacht. Aan de orde komen het leren van begrippen, algoritmen, probleemoplossen en het 'leren leren'.

Met name in de twee laatste onderwerpen is recente literatuur verwerkt, zoals de bekende studie van Scardamalia en Bereiter inzake het leren redigeren van teksten en van Palinscar en Brown over het rolwisselend onderwijzen bij het aanleren van leesstrategieën. De nieuwste ontwikkelingen uit de Verenigde Staten zijn echter nog niet verwerkt. We duiden hierbij op projecten waarin veelvuldig gebruik wordt gemaakt van video-disks de Jasper series en CSILE (een soort interactieve tekstdatabank waarmee leerlingen communiceren). Deze worden echter wel kort aangehaald in hoofdstuk 7. Het deel over

sociaal-affectief leren en psychomotorisch leren valt een stuk bescheidener uit, hetgeen het hoofdstuk wat onevenwichtig maakt.

Een centrale boodschap is dat de rol van de leraar bestaat uit het creëren van een krachtige leeromgeving. Verschaffel onderscheidt daarbij 5 kenmerken: een samenhangende aanpak, waarin aandacht is voor inhoud en strategie, 2) geleidelijke overdracht van docentsturing naar leerlingsturing, 3) leren in reële contexten, waarbij aandacht besteed wordt aan transfer, 4) werken in (kleine) groepen, 5) aandacht voor een positief leerklimaat.

Hoofdstuk 6 (Nijhof) handelt over de realisering van het curriculum. Hierbij worden het macro-, meso- en microniveau beschreven. Onder het macro- en mesoniveau worden opvattingen over het curriculum beschreven; het als achterhaald beschouwde vakkencurriculum, het curriculum als leerervaring met het gevaar dat doel en middelen los komen te staan van elkaar, het op competentie gerichte curriculum met een sterke doeloriëntatie, het curriculum als gelegenheid om te leren.

Vervolgens wordt de invloed van scholingsconcepten op het curriculum beschreven. Nijhof illustreert aan de hand van de basisvorming hoe moeizaam op landelijk niveau curriculaire keuzen tot stand komen, vaak vanuit tegengestelde basisoriëntaties (bijvoorbeeld de sociale adaptatie benadering versus de zelfrealisatie benadering). Hoe kunnen we dan van leraren verwachten dat zij helderheid hebben over de manier waarop zij curricula moeten gebruiken?

Het deel over de leraar als gebruiker van curricula valt tegen. Waar men het perspectief van de leraar voorstaat, zou men meer verwachten over het feitelijk gebruik van curricula, of beter gezegd leerboeken in de klas. Hierover wordt summier gerapporteerd, terwijl zeker meer bekend is over dit onderwerp, zoals bijvoorbeeld onderzoek op het CLU te Utrecht of buitenlands onderzoek naar onderwijsstijl en leermiddelen. Storend is de fout op p. 202. Niet de Centrale Registratie Leermiddelen maar het NICL (al enige jaren zelfs!) als afdeling binnen de SLO voert het beheer over de beschrijving van leermiddelen.

In hoofdstuk 7 beschrijft Lowyck hoe de leraar, onder invloed van een aantal factoren als leraarkekenmerken, doelstellingen, leerprocessen, interactiepatronen, randvoorwaarden en leerling-

kenmerken komt tot het inzetten van didactische werkvormen en media. De werkvormen zijn, geheel in de geest van het boek, gerangschikt naar de mate van sturing die de leraar of leerling heeft bij het ontstane leerproces. Het ontbreken van andere – meer didactische – ordeningscriteria, zoals naar doelen en organisatievorm (klas, groep, individu) illustreert de omslag in het denken in de richting van beïnvloeden van leerprocessen. In didactische handboeken zijn dergelijke ordeningen nog volop in zwang. Lowyck biedt in zijn beschrijving een ruim inzicht in resultaten van klassiek didactisch onderzoek, waarin met name aandacht is voor het gebruik van werkvormen, de typen vragen, het soort opdrachten, de soort interactie tussen leraren en leerlingen. Dit wordt gecombineerd met een beschrijving van de gevolgen voor het leren van leerlingen.

De overduidelijke boodschap in het onderdeel media is dat het bij het inzetten hiervan gaat om de vraag welke verwerkingsactiviteiten bij de leerlingen gerealiseerd kunnen worden.

Hoofdstuk 8, Didactische evaluatie door Verloop en Van der Schoot, gaat in op het belang van didactische evaluatie als hulpmiddel bij het nemen van beslissingen waarvoor de docent of het docententeam zich gesteld ziet. Evaluatie wordt in drie typen beslissingen uiteengezet:

a) plaatsing, selectie en resultaatbepaling, b) diagnose en c) sturing van het onderwijsleerproces. De cruciale rol van de docent in het evaluatieproces wordt benadrukt. Twee hoofdoorzaken van het toenemende belang van evaluatie worden aangegeven. Ten eerste omdat er een toenemende nadruk te zien is op kwaliteitszorg en kwaliteitsbepaling en ten tweede omdat in alle geledingen van het onderwijs een toenemende aandacht te zien is voor 'individuele leerwegen'. De computer zal in de toekomst tal van mogelijkheden bieden om evaluaties toe te spitzen op de individuele leerweg van de leerling. Hierdoor kan de computer een hulpmiddel vormen voor de ondersteuning van het onderwijsleerproces.

In hoofdstuk 9, Integrale kwaliteitszorg door Hoeben, wordt de kwaliteit van het onderwijs afgemeten aan de toegevoegde waarde die het onderwijs realiseert aan kennis en vaardigheden van de leerling. Het handelen van de docent wordt hierbij als zeer belangrijk gezien. Het schema voor integrale kwaliteitszorg wordt op

drie niveau's: leraar, school en overheid, behandeld. Opvallend is dat afzonderlijke behandeling of terugverwijzingen naar de onderwerpen 'effectieve instructie' en 'leerlingvolgsystemen' (Hst. 8) bij het hoofdstuk over integrale kwaliteitszorg mist.

Onderwijskunde is over het algemeen een vlot leesbaar boek, hoewel het niveau voor de PABO-studenten mogelijk wat te hoog zal liggen. Als naslagwerk voor onderwijsonderzoekers, onderwijswetenschappers (in opleiding) en onderwijsverzorgers biedt het vele ingangen om verder te studeren of inspiratie op te doen voor de eigen beroepspraktijk. Gebruik wordt gemaakt van een omvangrijk notenapparaat aan de hand waarvan de lezer zich verder kan verdiepen in de materie.

Als inleidend leer- of studieboek vraagt het veel van studenten. In een kort bestek komen tamelijk veel verschillende onderwerpen aan bod, die vragen om nadere concretisering en integratie. Voor de Nederlandse opleidingssituatie vult dit boek duidelijk een 'gat in de markt' met een goede weergave van de recente stand van zaken binnen de onderwijskunde.

G. Kanselaar en E. Roelofs

W. C. de Geus

Op zoek naar functionele leesvaardigheid

Academisch proefschrift. Vrije Universiteit Amsterdam
1995 228 pagina's ISBN 90 723 5337 4

'Schriftelijke zelfredzaamheid – kent U die uitdrukking?' zou dominee Gremdaat zeggen.

'Schriftelijke zelfredzaamheid' is de mooiste en meest kernachtige omschrijving van het begrip 'functionele geletterdheid' die ik in de dissertatie van Wendy de Geus tegenkwam. In deze dissertatie wordt een zoektocht beschreven naar een belangrijk aspect van functionele geletterdheid: functionele leesvaardigheid. Theoretisch wordt functionele leesvaardigheid onderscheiden van schoolse leesvaardigheid en de zoektocht is erop gericht dit onderscheid ook empirisch aan te tonen. Wanneer dit onderscheid duidelijk aanwezig is, zou het de moeite waard zijn om in het onderwijs aan functionele leesvaardigheid apart aandacht te besteden.

In hoofdstuk 1 wordt het begrip functionele lees-

vaardigheid gedefinieerd: het "... verwijst naar de vaardigheden die nodig zijn voor het zelfstandig en doelmatig kunnen lezen van voor het functioneren in Nederland belangrijke schriftelijke materialen" (p. 14).

Omdat de hoofdvraag van het beschreven onderzoek betrekking heeft op verschillen tussen functionele en schoolse leesvaardigheid, wordt ook schoolse leesvaardigheid gedefinieerd. He-las echter wordt het onderscheid met schoolse leesvaardigheid naar mijn mening niet vol-doende duidelijk gemaakt. De Geus zegt dat het vooral gaat om verschillen in taakdomeinen, dat de "schoolse leesvaardigheid nodig is voor het uitvoeren van schoolse leestaken en functionele leesvaardigheid voor het uitvoeren van allerlei praktische of alledaagse leestaken" (p. 11). Hier-mee wordt het definitieprobleem verschoven naar de aard van de leestaken. Een niet ge-noemd, maar wel essentieel verschil tussen schools en functioneel lezen is mijns inziens het volgende: Lezen op school doe je omdat het moet, het lezen buiten school wordt gedaan om-dat je dat zelf nodig vind. Een verschil in moti-vatie dus, hoewel je op school ook kunt lezen omdat je dat zelf nodig vindt (om een hoog cijfer te halen: dus functioneel) en je buiten school ook wel eens wat leest omdat het moet.

Ook in de manier waarop in de beschreven onderzoeken functionele en schoolse leesvaar-digheid zijn gemeten, is het onderscheid tussen die twee niet duidelijk. Functionele leesvaardig-heid werd gemeten met open vragen over een in-formatieve tekst, schoolse leesvaardigheid door middel van een tekst met meerkeuzevragen. Een probleem daarbij ligt in de ook hiervoor al ge-noemde motivatie om iets te lezen. In de praktijk zoek je informatie op vanwege een duidelijke noodzaak of vanuit een intrinsieke behoefte. In dit onderzoek daarentegen werd informatie op-gezocht omdat de onderzoeker dat zo graag wilde. In dit opzicht verschilden de schoolse en de functionele leestaken niet erg van elkaar. Voor de functionele leestaak wordt door de auteur een en ander ook onderkend in bijlage 2: "Zo zijn de taken niet echt functioneel..." (p. 191).

Het tweede probleem betreft de responsvorm bij het schoolse lezen. Het lezen van een tekst en vervolgens de alternatieven om daaruit het juiste antwoord te kunnen selecteren, lijkt mij eigenlijk een *functionele* leestaak. Aan de andere kant is

bij de gekozen functionele leestaak de respons-vorm verbaal. In de 'functionele werkelijkheid' is de respons op functioneel lezen vaak ook van andere aard: bijvoorbeeld een telefoonnummer draaien, een vertrektijd noteren, een gebruiks-aanwijzing opvolgen, etc.

Is overigens het onderscheid functioneel/schools alleen bij schoolleerlingen te maken of ook bij volwassenen?

Hoe functionele leesvaardigheid kan worden ge-meten wordt in hoofdstuk 2 behandeld. Daarbij worden verschillende methoden besproken in het licht van psychometrische criteria. Van de be-sproken meetmethoden lijkt de functionele lees-taak de beste. Hiermee wordt een *niveau* van functioneel lezen bepaald, hetgeen de voorkeur heeft boven een dichotome score van wel of niet functioneel leesvaardig. In hoofdstuk 3 worden diverse soorten functionele leestaken besproken.

De hoofdstukken 4 en 5 beschrijven respectieve-lijk twee uitgevoerde studies naar het verschil tussen functioneel en schools lezen. In beide on-derzoeken zijn de criteriumvariabelen de hier-voor reeds genoemde tekstbegrijptests voor het schoolse lezen en een functionele leestest voor het functioneel lezen.

De voorspellers in beide onderzoeken betref-fen vrij algemene cognitieve tests: kennistests en redeneertests. In de tweede studie is een woor-denschattest en een vragenlijst naar leesactivitei-ten toegevoegd aan het arsenaal van voorspel-lers.

In de eerste studie worden de steekproeven gevormd door leerlingen uit groep 8 van het ba-sisonderwijs ($n=218$) en leerlingen uit het eind van het derde leerjaar van het voortgezet onder-wijs ($n=279$). Van de 38 aangeschreven basis-scholen deden er elf mee, van de 25 aangeschre-ven VO-scholen deden er acht mee. De steekproefbeschrijving is wel heel erg summier. Niet vermeld is het selectie criterium op grond waarvan de scholen zijn aangeschreven, ook niet vermeld is in hoeverre de weigeringsgronden van de niet-meewerkende scholen samenhangen met de onderzoeksvariabelen. De afnames van de tests waren klassikaal en werden door de leer-krachten uitgevoerd, die daartoe een schriftelijk instructiepakket kregen. Nergens wordt gespro-ken over controle op de standaardisatie van de afnames. Al met al lijkt het enige sterke punt het

redelijke proefpersonen-aantal te zijn. Dat dan vervolgens met vertoon van ingewikkelde statistische en psychometrische termen geavanceerde analyses op de data worden beschreven, staat volgens mij in geen verhouding tot de kwaliteit van de data.

Dat de datakwaliteit in het eerste onderzoek te wensen over laat, is voor De Geus een reden om een verbeterd onderzoek op te zetten, welk replicatieonderzoek in hoofdstuk 5 wordt beschreven. De standaardisatie van en controle op de afnames is in deze tweede studie is ten opzichte van de eerste sterk verbeterd, door bij de afnames proefleiders aanwezig te laten zijn.

De steekproefbeschrijving is zo mogelijk nog summierder dan die bij het eerste onderzoek. Vermeld is wel hoeveel basisschoolgroepen (12 groepen; $n=301$) en VO-klassen (13 klassen; $n=291$) van de verschillende schooltypen participeerden, maar niet hoe die groepen geselecteerd zijn. Welke scholen zijn aangeschreven en waarom, hoeveel er zijn aangeschreven, hoe hoog het responspercentage is, of en waarom medewerking geweigerd is, is niet vermeld. Waarom dit belangrijk is, is dat naar mijn ervaring met testijks-onderzoek veel scholen in stedelijke gebieden met veel allochtone leerlingen hun medewerking vaker weigeren dan andere scholen. En, zoals we op pagina 13 van het proefschrift kunnen lezen, is functionele geletterdheid gebonden aan de sociale context: sociale en culturele omgeving.

Om de representativiteit ten aanzien van schooltype te schatten, wordt de VO-steekproefverdeling vergeleken met een landelijke verdeling van vierdeklassers. Hieruit wordt geconcludeerd dat de representativiteit 'redelijk' is. Een chi-kwadraattoets op tabel 5.1, over de vier vertegenwoordigde VO-niveaus, geeft echter een significant verschil te zien met de landelijke verdeling: $\chi^2=74.87$ ($df=3$, $p<0.01$).

Hoewel het tweede onderzoek een verbetering is ten opzichte van het eerste onderzoek, moet toch opnieuw geconcludeerd worden dat de kwaliteit van de steekproeven matig is.

De uitgevoerde analyses in de hoofdstukken 4 en 5 betreffen voornamelijk de toepassing van LISREL. Met LISREL kunnen modellen over de samenhang tussen variabelen getoetst worden. Naar mijn idee is LISREL hier echter zodanig

toegepast, dat er sprake is van een exploratieve analyse. Er worden zoveel modellen met elkaar vergeleken (acht in hoofdstuk 4; tien in hoofdstuk 5), dat er sprake moet zijn van kanskapitalisatie, zeker als de keuze voor een model berust op kleine verschillen in de toetsresultaten, zoals op pagina 105. In hoofdstuk 4 blijken bij de basisschoolkinderen drie modellen een lang niet significante chi-kwadraat te hebben en dus te passen. Bij het VO blijkt geen van de tien modellen te passen: alle chi-kwadraten zijn significant. Toch wordt het minst significante model (p toch altijd nog maar 0.001) serieus genomen en verder uitgewerkt.

Hoewel in de basisschoolgroep in hoofdstuk 5 alle tien modellen een significante chi-kwadraat opleveren, kunnen we toch in de tekst lezen dat alle modellen ongeveer hetzelfde *passen*. Er wordt toegegeven dat de chi-kwadraten significant zijn, maar dan wordt als een *deus ex machina* plotseling een parameter ('mse') ten tonele gevoerd, die, omdat deze kleiner is dan 0.08, zou wijzen op een 'redelijke benadering van het model dat in de populatie houdbaar zal zijn' (p. 135). Het zal best, maar verwerpings/bevestigingscriteria behoren vooraf te worden vastgesteld.

Gezien het toch exploratieve karakter van de wijze waarop LISREL is toegepast, had volgens mij evengoed gewone hoofdcomponentenanalyse gevolgd door rotatie toegepast kunnen worden. En waar ik al lezend naar snakte, waren tabellen met gewone produktmomentcorrelaties tussen de manifeste variabelen.

De uitkomsten van de twee onderzoeken zijn niet eenduidig weer te geven. In het eerste onderzoek kon niet worden aangetoond dat aan functioneel en schools lezen verschillende determinanten ten grondslag liggen, in het tweede onderzoek kon dat wel, zij het gedeeltelijk. Op het BO wordt het schoolse lezen voorspeld door woordenschat en de frequentie waarmee een leesboek gelezen wordt; het functionele lezen door kennis over teksten. Op het VO wordt het schoolse lezen verklaard vanuit redeneervaardigheid, algemene kennis en de frequentie waarmee een leesboek wordt gelezen; het functionele lezen slechts door verbaal redeneren (p. 142).

In hoofdstuk 6 wordt verslag gedaan van een onderzoek naar de determinanten van de moeilijkheid van functionele leesitems. Nagegaan werd

in hoeverre de p-waarden (van kinderen uit groep 8) van de items voorspeld konden worden vanuit een aantal taakkenmerken. Het resultaat is dat ongeveer de helft van de variantie kon worden voorspeld door een in de verschillende analyses wisselende combinatie van drie kenmerken. Het percentage verklaarde variantie is aan de lage kant en de uitkomsten zijn niet consistent. Een verklaring daarvoor is wellicht dat de p-waardenspreiding bij basisschoolleerlingen anders is dan bij bijvoorbeeld volwassenen. De Geus vermeldt in dit verband een correlatie van .82 met een set p-waarden bepaald bij 15-jarigen (waarom deze niet ook als criterium genomen?). Ik waag mij aan nog een andere mogelijke verklaring: misschien zijn er bij de verwerking van de informatie uit de teksten wel *typen* informatieverwerkers te onderscheiden, waarbij bij het ene type andere taakkenmerken een rol spelen dan bij een ander type.

Een analyse van taakkenmerken is niet uitgevoerd bij schoolse leestaken. Dit is jammer, omdat nu niet duidelijk is of de resultaten specifiek zijn voor functionele leestaken. Het lijkt me dat didactische aanbevelingen voor functioneel lezen meer op basis van dit soort analyses gedaan kan worden dan op basis van de in de hoofdstukken 4 en 5 gerapporteerde analyses, v.a. in vooral cognitieve vaardigheden van leerlingen die voorspellende variabelen zijn.

Concluderend kan worden opgemerkt dat de kwaliteit van de beschreven empirische onderzoeken moeilijk te beoordelen is omdat de steekproeven veel te summier zijn beschreven. Ondanks de kritiek moet toch worden opgemerkt dat De Geus een boeiend onderwerp heeft beschreven. De verdiensten van het proefschrift liggen naar mijn mening vooral in de uitvoerige conceptuele analyses van het begrip functionele leesvaardigheid. Helaas komen verschillen tussen functioneel en schools lezen qua definitie niet goed uit de verf, maar dit is gemakkelijker te signaleren dan te verhelpen, gezien de complexe aard van de materie.

H. C. Lutje Spelberg

Mededelingen

Schalen op onderwijskundig terrein

Door SVO is in 1994 een subsidie toegekend voor een project met als doelstelling: „onderzoek naar de mogelijkheden tot het samenvatten, vergelijken analyseren en systematiseren van informatie over in onderwijskundig onderzoek gebruikte instrumenten en het beschikbaar stellen daarvan via een modern medium.”¹

De achtergrondgedachte bij dit project was dat veel onderwijsresearchers zich op enig moment geconfronteerd zien met de vraag naar een bepaald meetinstrument. Om te helpen voorkomen dat het wiel steeds weer opnieuw wordt uitgevonden, zijn in 1994/1995 110 onderwijskundige schalen verzameld, waarbij zij opgemerkt dat het gaat om in enquête-onderzoek gebruikte schalen, niet om in individueel-diagnostisch onderzoek gebruikte tests.

Voor het zoeken van schalen is voornamelijk gebruik gemaakt van het PROMOND-bestand van het SWIDOC² dat een zeer uitgebreide verzameling proefschriften bevat op het gebied van mens- en maatschappijwetenschappen. Dit bestand kan lokaal geraadpleegd worden. Een probleem is dat het PROMOND-bestand niet ontsloten is op gehanteerde methoden en technieken en dus zeker niet op 'schalen'. Dit betekende dat de proefschriften ter hand moesten worden genomen. Het zoeken naar 'schalen' in de proefschriften viel tegen en kostte veel tijd omdat:

- in de inhoudsopgave zelden iets over een concrete methode staat genoteerd, hooguit een paragraaf 'methoden of instrumentconstructie';
- vaak de inhoudelijke argumentatie, de concrete operationalisatie, de psychometrische gegevens, de resultaten met de vragenlijst en de vragenlijst zelf op verschillende plaatsen vermeld werden;
- soms de vragenlijst (schaal) niet in zijn geheel werd weergegeven en slechts enkele items als voorbeeld vermeld werden; soms ook werden de items in verkorte vorm weergegeven;

- soms de antwoordalternatieven niet vermeld of slechts globaal vermeld werden;
- soms wel de antwoordalternatieven vermeld werden maar niet de schaalwaarden;
- soms het in de tekst vermelde aantal items niet met het aantal items in de bijlagen correspondeerde;
- soms de nummering van de items in de tekst een andere was dan die in de bijlagen;
- soms een schaal slechts kort genoemd werd en voor details door verwezen werd naar een intern rapport of een SVO-rapport.

Het zal na deze opsomming duidelijk zijn dat het aanbeveling verdient om alle relevante gegevens over een schaal of subschaal te vermelden op één plaats.

Per schaal zijn steeds de volgende rubrieken gehanteerd: naam schaal/ auteurs/ object/ respondenten (doelgroepen)/ sleutelverwijzing/ typering schaal/ aantal items/ aantal antwoordcategorieën/ psychometrische gegevens/ overige informatie/ instructie/ schaalitems/ commentaar.

Uit het overzicht van de rubrieken per schaal blijkt dat de meeste algemene gegevens (naam schaal, naam auteur, doelgroepen, bron, etc.) geen toelichting behoeven. Voor enkele punten is dit wel het geval. Achtereenvolgens geven we een toelichting op

a) *schaalnamen*, b) *object* en op b) *typering schaal*.

ad.a Schaalnamen

Opvallend is dat schaalnamen soms nietszeggend zijn: 'belangstelling, perspectief, motivatie, inzet' worden vaak gelijk gesteld aan 'attitude'. Het risico van een 'onjuiste naamgeving' is dat schalen onontdekt kunnen blijven door onderzoekers omdat de schalen onder een totaal vreemde titel een eigen leven leiden in een boek dat niet op het vakgebied van de onderzoeker ligt. Vandaar dat de naam die aan een schaal gegeven is, terwille van de herkenbaarheid soms iets uitgebreider is dan de naam die de auteur aan de schaal gaf.

ad.b Object

Het object legt het inhoudelijk domein vast. Dit domein heeft betrekking op

1. een eigenschap van de respondent zelf, of

2. een eigenschap van groepen waarvan de respondent zelf deel uitmaakt, of
3. een eigenschap van groepen waarvan de respondent niet zelf deel uitmaakt, of
4. een eigenschap van meer abstracte categorieën.

Zo kan aan individuele leerlingen (de *respondenten*, die we *doelgroep* noemen) een schaal voorgelegd worden met betrekking tot de volgende objecten:

- a de eigen inschatting van zijn/haar prestaties;
- b de mening over de inzet van 'de leerlingen in de klas';
- c de mening over de hulpvaardigheid van 'de docenten';
- d de waardering van een bepaald vak.

Bij 'object' wordt een primaire en een secundaire indeling gehanteerd, gescheiden door een "/".

De *primaire indeling* betreft een aantal 'sociale' categorieën (school; schoolleiding; schoolteam; docenten; leerlingen; leerlingbegeleiders; schoolbestuur; oudercommissies; zwakbegaafden; adolescenten) en enkele rechtstreeks op het onderwijs betrekking hebbende categorieën (onderwijs; vakgebied; studieonderdeel en nascholing).

De *secundaire indeling* brengt waar dit zinvol is een onderverdeling aan binnen elk van de primaire categorieën.

Indien aangegeven staat: 'docenten / persoonlijke verhouding met leerlingen' betekent dit dat aan respondenten (dat kunnen individuele docenten of leerlingen of wie dan ook zijn; een en ander blijkt uit de omschrijving van de doelgroep) is gevraagd wat hun mening is over hoe "docenten in het algemeen op hun school met leerlingen persoonlijk omgaan".

Het gaat hier dus niet om de meting van een eigenschap, zoals een attitude, van de respondent zelf, maar de respondent treedt op als informant over een sociale groep of systeem. Het domein van onderzoek is daarbij dus niet het individuele gedrag, maar een groter systeem, waarbij bij voorbeeld gemiddelde en spreiding van de percepties van respondenten over eigenschappen van dat systeem gemeten worden.

b) *Typering schaal*

De volgende indeling is gebruikt:

Likertschaal
Guttman/Mokkenschaa
Raschschaa
Semantische Contrastparen
Checklist

Toelichting bij 'Checklists'

De items hebben betrekking op los van elkaar staande elementen van een objectdomein die geenszins noodzakelijkerwijs met elkaar moeten correleren. De kwaliteit wordt dus exclusief bepaald door de volledigheid of de representativiteit van de gekozen elementen. Bij een checklist is het dus niet nodig om de betrouwbaarheid (Cronbachs alfa) te berekenen; slechts vanuit een oogpunt van exploratie kan dit zinvol zijn. Een groot aantal van de in deze documentatie opgenomen 'schalen' kan getypeerd worden als 'checklist', en aan de door de auteurs veelal berekende en door ons vermelde alfa's behoeft weinig betekenis te worden toegekend.

De verzamelde schalen worden thans aan geïnteresseerde gebruikers ter beschikking gesteld via Internet, op adres: <http://www.educ.uva.nl/schalen>. Daarnaast kan men eventueel bij de auteur op onderstaand adres een diskette aanvragen dat hetzelfde bestand bevat.

Het is voorts de bedoeling dat het schalenbestand regelmatig zal worden bijgewerkt, dat wil zeggen dat aanvullingen zullen worden opgenomen en dat fouten zullen worden hersteld. Op en aanmerkingen over de opgenomen schalen worden dus op prijs gesteld, evenals informatie over nieuwe, oude schalen die in aanmerking komen om in het schalenbestand te worden opgenomen. De zoektocht naar nieuwe, oude schalen gaat dus verder.

A. A. J. van Peet

tel. receptie: 020 525 1201; tel. A. v. Peet: 020 525 1263; E-mail: arie@educ.uva.nl; fax: 020 525 1200.

Noten

1. Het project stond onder leiding van prof.dr. P. G. Swanborn en dr. J.J. Hox en werd uitgevoerd door dr.A.A. v. Peet, allen verbonden aan de UvA, faculteit POW.
2. SWIDOC, Herengracht 410, 1017 Amsterdam. tel. 020 6225061, fax 020 6238374.

Congres Mediatheken en Bibliotheken

Op 28 mei 1996 vindt in Amersfoort het congres 'Mediatheken en Bibliotheken' - opzet, inrichting en integratie in het studiehuis' plaats. Dit congres dat georganiseerd wordt door de Zeister Openbare Bibliotheek, in samenwerking met de Stuurgroep Profiel Tweede Fase, biedt een landelijk platform voor allen die zich in het onderwijs en bibliotheekwezen verantwoordelijk voelen voor dit belangrijke studiehuisonderdeel. Tijdens dit congres zal het eindrapport van het inventarisatie-onderzoek naar schoolbibliotheken gepresenteerd worden.

Voor meer informatie: MesoConsult, Goudnodlaan 15, 5059 AE Tilburg; tel.: 013 4560311/4563276.

Congres Dyslexie

Het Dyslexie Congres 1996, georganiseerd door de afdeling Orthopedagogiek van de Rijksuniversiteit Groningen en de afdeling KNO van het Academisch Ziekenhuis te Groningen, in samenwerking met de landelijke Stichting Dyslexie Nederland, wordt op 5 juni te Groningen gehouden.

Voor nadere informatie en inschrijfformulieren: Congres Dyslexie 1996, Postbus 1536, 9701 BM Groningen; tel.: 050 3636584 (van 12.00 tot 15.00 uur); fax: 050 3636521 t.a.v. J. Knot, Dyslexie '96; E-mail J. Knot @ppsw.rug.nl

Oratie

Dr. L. T. Verhoeven, benoemd tot hoogleraar in de Orthopedagogiek aan de Faculteit der Sociale Wetenschappen van de Katholieke Universiteit Nijmegen zal op 31 mei 1996 zijn ambt aanvaarden met het uitspreken van een rede getiteld: *Omgaan met verscheidenheid in leren.*

Plaats: Aula Maior van het Aula/
Congresgebouw van de Katholieke Universiteit,
Comeniuslaan 2, Nijmegen
Tijd: 15.45 uur

Promoties

Universiteit van Amsterdam
Promovendus: A. S. Sabattini
Titel proefschrift: *Streetchildren. On becoming a streetchild*
Promotor: Prof.dr. M. G. M. van den Dungen
Datum: 10 juni 1996
Tijd: 13.30 uur

Universiteit van Amsterdam
Promovendus: E. M. M. Bros
Titel proefschrift: *School- en beroepsloopbanen van een Enschedese generatie*
Promotor: Prof.dr. J. Dronkers
Co-promotor: Prof.dr. A. Wesselingh
Datum: 27 september 1996
Tijd: 11.30 uur

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
21e jaargang, nr. 1, 1996

Onderwijsbeleid: ondersteunend of onrust
zaaiend? door R. Vandenbergh
Morele psychologie van deugd en zorg. Aan
gene zijde van orthodoxe ethiek, door S. E. Cuy-
pers
Care and feminist ethics. Noddings, Ruddick.
Baier, Benhabib, door L. Prior McCarthy
Over 'vertrouwen' en 'zorg' als morele en als
pedagogische basisconcepten, door P. Smeyers
Besprekingsartikel: Een feministische visie op
ethiek en pedagogiek: Nel Noddings' Ethic of
Care, door P. Smeyers

Tijdschrift voor Orthopedagogiek
35e jaargang, nr. 2, 1996

Themanummer Spelpedagogische problemen
Interspel en speelgoed. Redactionele inleiding,
door R. de Groot en E. J. J. de Bruyn
Spel en speelgoed in deze tijd, door R. de Groot,
J. M. Blom en C. J. M. Specken
De neiging tot spelen. Exploratief onderzoek
naar speelsheid als een zich ontwikkelend per-

soonlijkeidskenmerk, door E. J. J. de Bruyn, H. Rost, L. van der Poel en M. Riksen-Walraven
 Spelen en taal. Onderzoek naar relatie symbolisch spel, taalbegrip en taalproductie bij kinderen van twee en een half jaar, door W. J. M. Hamberg, W. P. Kunkeler en H. Rost
 Diagnostiek depressies jonge kinderen middels spelobservaties, door C. A. M. de Wit, A. Mol Lous, H. Rost en E. J. J. de Bruyn
 Spelopvoeding van kinderen met (lichamelijke) beperkingen, door R. de Groot

Tijdschrift voor Orthopedagogiek
 35e jaargang, nr. 3, 1996

Hoofdrekenen tot honderd op maat, door E. G. Harskamp en C. J. M. Suhre
 Dyslexie. Woordidentificatie-tests, fonologische gevoeligheid en symboolbenoemsnelheid, door K. P. van den Bos en H. C. Lutje Spelberg
 Integratie kinderen met Downs syndroom. Reactie op Scheepstra, Pijl en Nakken op Kraijer, door G. W. de Graaf en E. A. B. de Graaf

Kuhlemeier, J. B., *Taalvaardigheid, taalactiviteiten en taalattitudes. Een validatiestudie* (dissertatie). Cito, Arnhem, 1996.
 Raad voor het Jeugdbeleid, *Van verdrag tot gedrag. Rechten van kinderen getoetst in onderwijs en in jeugdhulpverlening en jeugdbescherming*. Uitgeverij SWP, Utrecht, 1996, f 19,50.
 Riet, N. van, & Wouters, H., *Casemanagement. Een leer-werkboek over de organisatie en coördinatie van zorg-, hulp- en dienstverlening*. Uitgeverij Van Gorcum, Assen, 1996, f 49,50.
 Schellings, G. L. M., *An educational approach to selecting main points in texts. Motives, methods, tasks and individual differences* (dissertatie). Drukkerij Foppele-Vinck, 's-Hertogenbosch, 1995.
 Zoontjens, F. J. J., Backx, H. A. M., & Mentink, D., *Kwaliteit en uitvoerbaarheid van onderwijswetgeving. Preadviezen Vereniging voor Onderwijsrecht*. Secretariaat van de VOR, Nassaulaan 6, 's-Gravenhage, 1995, f 29,75.

Inhoudsopgave

Pedagogische Studiën Jaargang 73, nummer 3/1996

S. Veenman, H. de Jonghe en E. van Wezel

Evaluatie van een trainingsprogramma voor het coachen van leraren 170

H. Blok en P. P. M. Leseman

Effecten van voorschoolse stimuleringsprogramma's: een review van reviews 184

B. Witziers

Coördinatie van het vakonderwijs binnen scholen voor voortgezet onderwijs: vaksecties en schoolleiding 198

J. W. Steutel

De deugdenbenadering van de morele opvoeding. Een poging tot conceptuele verheldering 215

Discussie

P. M. Kroonenberg en R. van der Veer

Internationale publikaties en kwaliteit. Een onderzoek naar het publikatiegedrag van Nederlandse pedagogen en onderwijskundigen 225

Redactie Pedagogische Studiën

De onderzoeksvisite 1995 239

Boekbesprekingen 242

Mededelingen 246

Evaluatie van een trainingsprogramma voor het coachen van leraren*

S. Veenman, H. de Jonghe en E. van Wezel**

Samenvatting

In dit onderzoek wordt verslag gedaan van de resultaten van een training in coachingsvaardigheden voor schoolbegeleiders. Coaching kan worden gezien als een van de manieren om leraren te voorzien van feedback op hun functioneren en hen te stimuleren tot zelfreflectie. Onderzocht werd of schoolbegeleiders de kennis en de vaardigheden uit het trainingsprogramma 'De schoolbegeleider als coach' toepasten in hun coachingsgesprekken met leraren. Daartoe werden vóór en na de training coachingsgesprekken geanalyseerd volgens een quasi-experimenteel 'pretest-posttest control group design'. Ook werd aan de gecoachte leraren gevraagd hun oordeel te geven over de toegepaste coachingsvaardigheden en over de mogelijke uitwerking van coaching op hun instructiegedrag. Een positief trainingseffect werd gevonden voor de coachingsvaardigheden die betrekking hadden op 'autonomie-bevordering', het 'geven van feedback' (tijdens het nagesprek) en op 'doelgerichtheid' ten aanzien van het formuleren van verbeteringsvoorstellen voor het instructiegedrag. Met betrekking tot het oordeel van de gecoachte leraren werden geen verschillen gevonden tussen gecoachte en niet-gecoachte leraren. Het effect van coaching op de verbetering van hun instructiegedrag werd door de leraren redelijk hoog ingeschat.

1 Achtergrond

Door het Christelijk Pedagogisch Studiecentrum (CPS) en de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen is een trainingsprogramma ontwikkeld voor schoolbegeleiders met betrekking tot het coachen van leraren. De evaluatiegegevens van het schoolverbeteringsprogramma 'Effectieve instructie en doelmatig klasmanagement' (Veenman, Lem, Roelofs, & Nijssen, 1994) vormden de aanleiding tot de ontwikkeling van dit coachingsprogramma. Uit de tweede evaluatiestudie bleek dat leraren die in aanvulling op het nascholingsprogramma coaching kregen op onderdelen van de gegeven nascholing meer profiteerden dan leraren die geen coaching ontvingen. Toch bleek het effect van coaching over het algemeen minder groot dan verwacht (Roelofs, Veenman, & Raemaekers, 1993). Een verklaring hiervoor was dat de schoolbegeleiders die de coaching verzorgden vooraf geen concrete handreikingen hadden gekregen over de uitvoering van coaching. Dit leidde er toe dat de schoolbegeleiders coaching op verschillende wijze gingen interpreteren. Zij bleken behoefte te hebben aan bruikbare richtlijnen en voorbeelden voor de praktijk. Derhalve werd besloten een trainingsprogramma voor coaching samen te stellen en te evalueren op bruikbaarheid. In dit voornemen werden we gesteund door het vrij sterke positieve verband dat Houtveen (1990) in haar onderzoek vond tussen het begeleiden van leerkrachten in de klas en het vernieuwend pedagogisch-didactisch handelen van de leerkrachten. Uit dit onderzoek kwam ook naar voren dat het begeleiden van leraren in de klas nauwelijks voorkomt. Het belang van het begeleiden van leerkrachten in de klas bleek eveneens uit de resultaten van het Landelijk Project Schoolverbetering. Een begeleidingsaanpak, bestaande uit een combinatie van coachingsge-

* Dit onderzoek werd mede mogelijk gemaakt door het Christelijk Pedagogisch Studiecentrum (CPS) te Hoevelaken.

** De auteurs bedanken de volgende personen voor hun bijdragen aan deze studie: K. Vernooy (CPS), C. Compagnie (CPS), Y. Leenders (CPS), L. Bouts (RTD KUN) en de deelnemende schoolbegeleiders en leraren.

sprekken van de leerkracht uit groep drie met de schoolbegeleider aan de hand van klasse-observaties, van begeleiding van het schoolteam als geheel en van begeleiding door de schoolleider bleek een positief effect te hebben op de taakgerichte leertijd tijdens leeslessen, op het instructiegedrag van de leerkracht, op de mate van planmatig werken en op de technische leesprestaties van de leerlingen (Houtveen, Booij & De Jong, 1995a,b).

Voor alle duidelijkheid moet worden opgemerkt dat het in dit artikel beschreven programma voor het coachen van leerkrachten géén onderdeel vormt van het schoolverbeteringsprogramma 'Effectieve instructie en doelmatig klasmanagement.' De hier beschreven resultaten hebben dan ook geen betrekking op dat programma. Het onderhavige trainingsprogramma voor het coachen van leerkrachten vormt een op zichzelf staand geheel.

2 Coaching

In het rapport 'Een beroep met perspectief: de toekomst van het leraarschap' van de Commissie Toekomst Leraarschap (1993) wordt opgemerkt dat het tot de professionaliteit van de leraren behoort om veranderingen en verbeteringen in het onderwijs in gang te zetten. De leraren worden beschouwd als de motor van de veranderingsprocessen. Om deze rol te kunnen vervullen moeten ze hiertoe gemotiveerd zijn en de instrumenten hiervoor aangereikt krijgen. De leraar als zelfstandige beroepsbeoefenaar die aan zijn lot wordt overgelaten strookt niet met de opvattingen over de verdere professionalisering van het leraarschap. De commissie merkt op: "Leraren hebben recht op feedback en beoordeling van hun professioneel handelen. Daarom moet het isolement van de beroepsuitoefening in de klas worden opgeheven. Een andere organisatie van het primair proces maakt het mogelijk om samen met collega's professionaliteit op te bouwen en de doelmatigheid van de taakverdeling te vergroten. Dat vraagt om professionele vaardigheden van leraren als zelfreflectie, in staat zijn taken te verdelen, met anderen overleggen, samen te werken, ook op het punt van lesgeven, en gewenste veranderingen in het eigen handelen door te voeren" (Commissie Toekomst Leraar-

schap, 1993, p. 76).

Coaching vormt een van de manieren om leraren te voorzien van feedback op hun eigen functioneren en hen te stimuleren tot zelfreflectie (Schön, 1983). Coaching is een begrip dat recent gebruikt wordt voor het op deze wijze professioneel ondersteunen van leraren. Het is nauw verwant aan het begrip 'klinische supervisie' dat een langere traditie kent. Hieronder wordt verstaan: "An in-class support system designed to deliver [...] assistance directly to the teacher [...] to bring about changes in classroom operation and teacher behavior" (Sergiovanni & Starrat, 1983, p. 299). Het woord 'klinisch' wil zeggen dat het gaat om een één-op-één-relatie tussen een supervisor en een leraar en dat het gaat om het feitelijke functioneren van de leraar: wat de leraar in de klas doet vormt het uitgangspunt voor het geven van supervisie. Kenmerkend hierbij is dat de supervisor via observaties in de klas gegevens verzamelt die het aangrijpingspunt vormen voor nadere analyse en reflectie. In de Nederlandse situatie wordt soms gesproken van 'klasseconsultatie' (Verkerk, 1987; Houtveen, Booij & De Jong, 1995a,b) of van 'directe consultatie' (Westerop, 1993).

Omdat bij supervisie vaak sprake is van een relatie tussen een bovengeschiedte en een ondergeschiedte en supervisie vaak gekoppeld is aan evaluatie of beoordeling, wat bij veel leraren negatieve emoties oproept, spreek men tegenwoordig liever van coaching. Dit begrip brengt beter tot uitdrukking dat het gaat om een helpende of ondersteunende relatie en niet om een beoordeling. Coaching wil zeggen dat men iemand wil helpen zich te ontwikkelen van een feitelijke naar een gewenste situatie. In het door ons ontwikkelde trainingsprogramma hebben we coaching omschreven als een vorm van gerichte ondersteuning van de leraar bij zijn werk in de klas met als doel zijn vermogen te versterken om zijn vakbekwaamheid blijvend te verbeteren en te vernieuwen op basis van een systematische reflectie op het professioneel handelen (Veenman, De Jonghe & Van Wezel, 1995). De op de professionele ontwikkeling van de leraar gerichte ondersteuning wordt gegeven door een coach en is gecentreerd rond de cyclus: voorbespreking, observatie en nabespreking. Voor een goed begrip in het coachingsproces is enig inzicht in de ver-

schillende modellen voor het begeleiden van leraren en hun historische achtergronden nodig. Bij de bespreking van deze modellen beperken we ons tot die modellen die ten grondslag hebben gelegen aan het trainingsprogramma 'Het coachen van leerkrachten'.

3 Enkele modellen voor klinische supervisie en coaching

Aan het eind van de jaren zestig legden Goldhammer en Cogan de grondslagen voor het model van klinische supervisie. Zij ontwikkelden dit model als een reactie op de traditionele vormen van supervisie met een sterk beoordelende inslag. De supervisor in de rol van expert maakte duidelijk wat de leerkrachten dienden te veranderen en hoe ze dat moesten doen. De supervisor gaf oplossingen voor problemen die in zijn ogen belangrijk waren, maar die niet noodzakelijkerwijs de problemen van de leerkrachten hoefden te zijn. De nadruk lag op de controle van het werk van leraren. In de modellen van Goldhammer en Cogan worden drie doelen nagestreefd: vertrouwen, leren en autonomie.

In het door Goldhammer (1969) ontwikkelde model voor klinische supervisie ligt de nadruk op wederzijdse collegialiteit en respect en op het gezamenlijk zoeken naar de zingeving van de gebeurtenissen in de klas. Het door Goldhammer geschetste model bestaat uit vijf fasen: (1) een voorbespreking waarin de doelen en de achtergronden van klinische supervisie worden belicht, (2) het uitvoeren van observatie(s) in de klas, (3) het analyseren van de verzamelde observatie-gegevens en het kiezen van de juiste strategieën voor de uitvoering van de nabespreking, (4) een nabespreking waarin de verzamelde gegevens worden geanalyseerd en (5) een achteraf-analyse van de nabespreking ter mogelijke verbetering van het supervisie-proces.

Anderson en Krajewski hebben het oorspronkelijke model van Goldhammer later geactualiseerd (Goldhammer, Anderson & Krajewski, 1980). Het basisstramien van het model van Goldhammer is hierbij gehandhaafd.

In het door Cogan (1973) uitgewerkte model ligt het accent op het verbeteren van het onderwijsgedrag van de leraren in de klas door middel van zelf-analyse en zelf-reflectie. Het doel van klinische supervisie is gericht op de ontwikkeling van een professioneel handelende leraar die in staat is zijn eigen functioneren te analyseren en die open staat voor de hulp aangeboden door anderen. Ook in dit model wordt gebruik gemaakt van de cyclus: voorgesprek, observatie in de klas en nagesprek.

De modellen van Goldhammer en Cogan weerspiegelen de uitgangspunten van de humanistische psychologie. De modellen van Hunter en Russell (1990) en Joyce en Showers (1988) hebben op grond van hun gedragswetenschappelijke achtergrond een meer technisch-didactische inslag (Pajak, 1993). De nadruk ligt in deze modellen op de verbetering van het onderwijsgedrag aan de hand van de uitkomsten met betrekking tot het onderzoek naar effectieve instructiestrategieën.

Voor Hunter en Russell (1990) ligt het doel van coaching in het versterken van het vermogen van de leraar om steeds professioneler beslissingen te nemen. Deze beslissingen dienen zoveel mogelijk gebaseerd te zijn op de uitkomsten van onderzoek over leren en instructie. Deze uitkomsten worden in de vorm van als-dan relaties aangeboden in de vorm van (na)scholing. Het model van Hunter en Russell bestaat uit de volgende vijf fasen: (1) scholing, (2) observatie, (3) analyse, (4) nabespreking en (5) follow-up. Volgens Hunter (1986) is een voorbespreking niet nodig omdat de leraren op grond van de nieuwe kennis en vaardigheden tijdens de scholingsfase zelf kunnen vaststellen op welke punten hun onderwijsgedrag voor verbetering in aanmerking komt. Bij het observeren in de klas wordt een beschrijvend verslag gemaakt van de gebeurtenissen in de klas.

Het model van Joyce en Showers (1982, 1988) is meer een model voor de vormgeving van effectieve scholing of training dan een begeleidingsmodel. Coaching van leraren wordt in dit model vooral gezien als een transferbevorderende maatregel (Bennett, 1987). Door Joyce en Showers (1980) wordt coaching omschreven als: "Hands-on, in-classroom assistance with the transfer and application of skills to the classroom".

Het model van 'cognitive coaching' van

Costa en Garmston (1994) steunt op de inzichten van de cognitieve psychologie. Bij cognitieve coaching wordt expliciet aandacht besteed aan de cognities of de mentale werkmodellen van leraren als middel om het onderwijsgedrag van leraren te verbeteren. Het werk van de coach is niet direct gericht op het veranderen van het uiterlijke onderwijsgedrag van de leraren. Dit gedrag verandert als gevolg van een herstructurering van cognitieve processen. Cognitieve coaching tracht drie doelen te realiseren: (1) het scheppen van een vertrouwensbasis tussen coach en leraar, (2) het stimuleren van wederzijds leren in de zin van het construeren van nieuwe inzichten en (3) het versterken van het autonoom handelen van de leraar. Het proces van cognitieve coaching bestaat uit een drietal fasen: een voorbespreking (planning), observatie en een nabespreking (reflecting).

Elk van de hier beschreven modellen bevat een aantal concrete supervisie- of coachingsvaardigheden. Ook worden richtlijnen gegeven voor het uitvoeren van observaties in de klas. Bij het samenstellen van een trainingsprogramma voor coaching is een selectie gemaakt van relevante elementen en coachingsvaardigheden uit deze modellen. Overgenomen is de coachingscyclus: voorbespreking, observatie, nabespreking. Voor een voorbespreking is gekozen omdat het trainingsprogramma niet direct gekoppeld is aan een scholingsprogramma zoals in de modellen van Hunter en Russell (1990) en Joyce en Showers (1988). Overgenomen zijn verder de drie doelen van supervisie of coaching: het scheppen van een vertrouwensbasis tussen leraar en coach, het stimuleren van het professioneel handelen van de leraar door middel van reflectie op het actueel handelen in de klas en het vergroten van de autonomie van de leraar door hem te stimuleren zelf met verbeteringsvoorstellen voor zijn handelen te komen. In het coachingsproces dient niet alleen aandacht geschonken te worden aan het actuele gedrag van de leraar maar ook aan de cognities die in verband staan met dit gedrag. De keuze van de coachingsvaardigheden is mede geïnspireerd door het trainingsprogramma *Another set of eyes* van de Association for Supervision and Curriculum Development (1988). In dit programma staan de volgende vaardigheden centraal: 'trust-building', 'ques-

tioning', 'responding' en 'empowerment'. Op grond van de opvattingen van Hunter en Russell (1990) is bij de observatie gekozen voor het beschrijvende verslag. Dit betekent dat de coach, op grond van de gemaakte afspraken in de voorbespreking van een les, een zo uitgebreid mogelijk protocol maakt en de observatie niet beperkt via voorgestructureerde observatieformulieren. Aldus wordt een informatieve beschrijving verkregen van de les die het uitgangspunt vormt voor de nabespreking. Op de inhoud en de structurering van het trainingsprogramma wordt verderop ingegaan.

4 Effecten van klinische supervisie en coaching

Naar de effecten van klinische supervisie en coaching is vooralsnog weinig empirisch onderzoek gedaan. Het meeste onderzoek is uitgevoerd in de Verenigde Staten. In hun analyse van zes case-studies met betrekking tot klinische supervisie trachtten Nolan, Hawkes en Francis (1993) het effect ervan na te gaan op de leraar. Zij stelden vast dat klinische supervisie onder de volgende condities leidt tot veranderingen in het denken en het gedrag van de leraar: (1) er dient sprake te zijn van een collegiale relatie waarbij de leraar zich veilig en ondersteund voelt, (2) de leraar dient de inhoudelijke kant van het proces zelf te kunnen sturen, (3) continuïteit van de ondersteuning dient verzekerd te zijn, (4) gerichte en nauwkeurige beschrijvingen van het feitelijke gedrag van de leraar en de leerlingen in de klas vormen het uitgangspunt voor reflectie en (5) reflectie door zowel de leraar als de supervisor vormt de kern van de nabespreking.

Studies waarin het effect van klinische supervisie op de leerprestaties van de leerlingen wordt onderzocht zijn schaars. Gezien het grote aantal interveniërende variabelen (zoals de kwaliteit van de uitgevoerde supervisie, de persoonlijkheidskenmerken van de supervisor en de supervisanten, de kenmerken van de school en de leerlingen, de kwaliteit van het instructieproces, de frequentie en de duur van het supervisieproces) is dit niet zo verwonderlijk. Na een uitgebreide speurtocht in de literatuur vond Pavan (1985) vier Engelstalige studies waarin getracht werd het effect van klinische

supervisie op de leerprestaties van de leerlingen vast te stellen. Van een duidelijk effect was geen sprake. Omdat deze studies vele methodologische tekortkomingen kenden, staat volgens Pavan het onderzoek naar het verband tussen klinische supervisie en de leerprestaties nog in de kinderschoenen.

Sparks en Bruder (1987) vonden dat het geven van coaching leidde tot een grotere collegiale samenwerking en tot een grotere openheid om te experimenteren met nieuwe instructievormen.

In de studies van Showers (1982), Baker (1983) en Bennett (1987) bleek coaching transfer van de trainingssituatie naar de werkplek te bevorderen. Leraren die naast nascholing ook coaching ontvingen pasten de nieuw geleerde instructievaardigheden vaker, adequater en langer toe dan niet-gecoachte leraren. Een effect op de prestaties van de leerlingen (vastgesteld via een essay-test) werd echter niet gevonden.

In een studie van Roelofs (1993) en Roelofs, Veenman en Raemaekers (1993) bleken gecoachte leraren een aantal vaardigheden op het gebied van effectieve instructie en doelmatig klasmanagement beter toe te passen dan niet-gecoachte leraren. Een verband tussen het geven van coaching en de leerprestaties van de leerlingen kon niet worden vastgesteld (Roelofs, Veenman, & Voeten, 1994). Met betrekking tot klasseconsultatie vond Snippe (1991) eveneens geen eenduidige effecten van deze begeleidingsaanpak op de rekenprestaties van de leerlingen. Wel bleek klasseconsultatie effect te hebben op de implementatie van een leergang cijferend rekenen. Zoals vermeld, constateerden Houtveen, Booij en De Jong (1995a,b) wel positieve effecten van coaching – in combinatie met andere vormen van begeleiding – op de technische leesprestaties van leerlingen uit groep drie.

Gezien de in de literatuur vermelde positieve effecten van klinische supervisie en coaching op de collegiale samenwerking en de openheid om te experimenteren met nieuwe instructievormen in een schoolteam, de versterking van het reflexieve vermogen van leraren als indicatie voor professionele groei, de overdracht van nieuw geleerde vaardigheden van de trainingssituatie naar de klaspraktijk, werd in overleg met het CPS besloten om voor school-

begeleiders een trainingsprogramma voor coaching te ontwikkelen.

5 Onderzoeksvragen

Dit onderzoek richt zich op de implementatie-effecten van het trainingsprogramma '*Het coachen van leerkrachten*'. Het onderzoek kent de volgende vraagstellingen:

- (1) Passen de schoolbegeleiders de kennis en vaardigheden uit het trainingsprogramma toe in de coachingsgesprekken die zij voeren met de leraren?
- (2) Leidt de gegeven coaching van de schoolbegeleiders volgens het oordeel van de leraren tot veranderingen in hun instructiegedrag?

6 Methode en instrumenten

6.1 Opzet van het onderzoek

Het onderzoek bestaat uit twee deelonderzoeken:

- (a) een evaluatie-onderzoek waarbij gebruik wordt gemaakt van de beoordelingen van getrainde beoordelaars en dat gericht is op de toepassing van de kennis en vaardigheden uit het trainingsprogramma door de schoolbegeleiders bij het voeren van coachingsgesprekken met leraren basisonderwijs;
- (b) een evaluatie-onderzoek waarbij gebruik gemaakt wordt van de beoordelingen van de leraren die door deze schoolbegeleiders zijn gecoacht.

De evaluatiestudie werd opgezet volgens een quasi-experimenteel 'pretest-posttest control'group design'. De experimentele groep in het eerste deelonderzoek bestond uit 18 getrainde schoolbegeleiders en de controlegroep uit 10 ongetrainde schoolbegeleiders. De experimentele groep uit het tweede deelonderzoek bestond uit 18 gecoachte leraren en de controlegroep uit 8 niet-gecoachte leraren.

6.2 De onderzoeksgroep

Door het CPS werd een brief gestuurd naar 60 schoolbegeleidingsdiensten in Nederland en één schoolbegeleidingsdienst in België met de vraag of schoolbegeleiders interesse hadden om deel te nemen aan het trainingsprogramma '*Het coachen van leerkrachten*'. Als voor-

waarde aan deelname werd gesteld dat de schoolbegeleiders geen of weinig ervaring moesten hebben met het coachen van leraren en in de nabije toekomst van plan waren leraren te gaan coachen. Verder moesten zij bereid zijn vóór en na de training coachingsgesprekken te voeren met een of twee leraren uit het basisonderwijs. Omdat het aantal aanmeldingen ($n = 39$) groter was dan verwacht, werd besloten aan alle geïnteresseerde schoolbegeleiders een training aan te bieden. Twee groepen werden gevormd. Groep één kreeg de training in september 1993, groep twee in januari 1994. Uit de totale groep van aanmeldingen werden 25 schoolbegeleiders op aselechte wijze toegewezen aan de eerste groep, zij vormden de experimentele groep. De 14 schoolbegeleiders uit de tweede groep die de training later volgden, vormden de controlegroep. Aldus bestonden zowel de experimentele groep en de controlegroep uit schoolbegeleiders die geïnteresseerd waren in coaching. Op deze wijze werd de uitgangssituatie van beide groepen gelijk gehouden en het effect van zelf-selectie beperkt.

Uiteindelijk namen 18 schoolbegeleiders uit de experimentele groep aan de training deel en 10 schoolbegeleiders participeerden in de controlegroep. De andere schoolbegeleiders konden onvoldoende tijd vrij maken of konden geen leraren vinden die bereid waren aan coachingsgesprekken deel te nemen.

De aan het onderzoek deelnemende schoolbegeleiders hadden een gemiddelde schoolbegeleiderservaring van 8 jaar en hadden over het algemeen weinig of geen ervaring met coaching. De gemiddelde leeftijd bedroeg 35 jaar. De groep bestond voor 67% uit vrouwelijke schoolbegeleiders en voor 33% uit mannelijke schoolbegeleiders.

Bij de 18 leraren die coaching ontvingen werden vervolgens gegevens verzameld over de wijze waarop de schoolbegeleiders de coaching uitvoerden en welke effecten zij van de coaching verwachten (response 100%). Dezelfde gegevens werden verzameld bij 8 leraren die coaching ontvingen van schoolbegeleiders die niet aan de training deelnamen (response 80%). De gemiddelde onderwijzerservaring van de aan het onderzoek deelnemende leraren bedroeg 12 jaar, de gemiddelde leeftijd was 36 jaar. De groep leraren bestond voor 70% uit vrouwen en voor 30% uit mannen.

6.3 De Schaal voor Coachingsvaardigheden

Om te kunnen vaststellen of de schoolbegeleiders na de training in staat zijn de geleerde coachingsvaardigheden in praktijk te brengen, werd de Schaal voor het Gebruik van Coachingsvaardigheden (SGCV) gebruikt. Dit instrument werd zowel gebruikt door de twee getrainde beoordelaars als de leraren die coaching ontvingen van de schoolbegeleiders. Bij de ontwikkeling van de schaal werd uitgegaan van vier dimensies, namelijk 1) het vermogen van de coach om de leraar te stimuleren om zelf met verbeteringsvoorstellen voor zijn instructiegedrag te komen, 2) de vaardigheid van de coach met betrekking tot het geven van feedback aan de leraar, 3) het vermogen van de coach het coachingsgesprek doelgericht te voeren, en 4) het vermogen van de coach het vertrouwen van de leraar te winnen. Voor elke dimensie werden een aantal vragen geformuleerd. Gebruik werd gemaakt van een vijfpuntsschaal (met als schaalwaarden: 1 = in het geheel niet, 5 = in zeer ruime mate). De oorspronkelijke schaal bevatte 32 items. Een aantal van de items werd ontleend aan de 'Snyder-Pavan Clinical Supervision Practices Questionnaire' (Pavan, 1991, 1993). De beoordeling van het coachingsgedrag van de schoolbegeleiders vond plaats aan de hand van op audiocassette vastgelegde coachingsgesprekken tussen schoolbegeleiders en leraren. De tussen-beoordelaar-betrouwbaarheid (geschat via Cohen's Kappa) varieerde voor de diverse items van 0.63 tot 1.00 (mediaan 0.85). Twee items werden uit de schaal verwijderd vanwege een te lage interbeoordelaar-betrouwbaarheid en twee items werden verwijderd vanwege een te geringe spreidingsbreedte. De uiteindelijke versie van de SGCV telde aldus 28 items. Via een principale-componenten-analyse werden voor de SGCV vier subschalen onderscheiden met als titel: (1) autonomie-bevordering (12 items, $\alpha = .94$), (2) feedback (5 items, $\alpha = .84$), (3) doelgerichtheid (7 items, $\alpha = .87$) en (4) het verwerven van vertrouwen (4 items, $\alpha = .61$). De totaal verklaarde variantie voor deze vier subschalen bedroeg 72%.

Voorbeelden van items uit de subschaal 'autonomie-bevordering' zijn: De coach stimuleert de leerkracht zelf met alternatieven te komen voor het verbeteren van zijn instructiegedrag; De coach vraagt de leerkracht aan te

geven wanneer hij van mening is dat zijn instructiegedrag is verbeterd. De subschaal 'feedback' bevat items als: De feedback van de coach is concreet en specifiek; De coach geeft beschrijvende feedback. In de subschaal 'doelgerichtheid' zijn items opgenomen als: De coach zorgt er voor dat de leerkracht terzake blijft; De coach blijft zelf gedurende het gesprek ter zake. De subschaal 'het verwerven van vertrouwen' bevat items als: De coach stelt de leerkracht op zijn gemak; De coach toont respect voor de leerkracht.

Bij de vragenlijst voor leraren werd nog een vijfde subschaal onderscheiden: de waargenomen effecten van het coachingsgesprek op het instructiegedrag (12 items, $\alpha = .88$). Voorbeelden van items uit deze schaal zijn: De verbeteringen die ik na het gesprek met de schoolbegeleider ten aanzien van mijn lesgedrag wil realiseren hebben een positieve uitwerking op: a) mijn functioneren in de klas, b) mijn omgang met leerlingen, c) het gebruik van methoden/lesmaterialen in de klas.

6.4 De uitvoering van het trainingsprogramma

Het werkboek 'Het coachen van leerkrachten' bestaat uit vijf hoofdstukken. In het eerste hoofdstuk (coaching en onderwijsverbetering) wordt ingegaan op de school als lerende organisatie en wordt gedefinieerd wat onder coaching wordt verstaan. Tevens worden de kenmerken en functie van coaching uiteengezet en de vaardigheden die de coach en de leraar nodig hebben om coaching tot een succes te maken. Om de coachingsvaardigheden te illustreren wordt gebruik gemaakt van voorbeelden die meestal betrekking op aspecten van het instructiegedrag van de leraar.

Coaching kan verschillende vormen aannemen (bijvoorbeeld 'counseling, mentoring, tutoring, confronting'). In het tweede hoofdstuk wordt uitgelegd dat gekozen is voor consulterend coachen. Hierbij wordt er van uitgegaan dat het de leraar is die het initiatief neemt om zijn functioneren in de klas kritisch te analyseren en ter verbetering van dit functioneren ondersteuning behoeft van een coach. Verder wordt in dit hoofdstuk de coachingscyclus uiteengezet (voorgesprek, observatie, nagesprek).

In het derde hoofdstuk worden de coachingsvaardigheden van het voorgesprek bespro-

ken. Aan de orde komen vaardigheden die betrekking hebben op het verwerven van vertrouwen, het afbakenen van een probleem, het oplossen van een probleem, het kiezen en uitwerken van een passend alternatief en het opstellen van een verbeteringsplan.

Het verzamelen van gegevens in de klas komt in het vierde hoofdstuk aan de orde. Ingegaan wordt op verschillende vormen van observatie en op het maken van een lesprotocol (het beschrijvende verslag).

In het vijfde hoofdstuk worden de vaardigheden van het nagesprek besproken. Aan de orde komen vaardigheden die betrekking hebben op 'stoom afblazen', op het evalueren van de oplossing, op het bespreken van de geobserveerde les (o.a. het geven van feedback en reflecteren), op het ontwikkelen van nieuwe alternatieven en op het bijstellen van het verbeteringsplan.

Aan zowel het eind van het voor- als nagesprek vindt een evaluatie van het uitgevoerde coachingsgesprek plaats. De onderscheiden coachingsvaardigheden van het voor- en nagesprek zijn gestructureerd rond drie fasen zoals onderscheiden in het werk van Kinlaw (1989): 'presenting, using reaction to develop information, resolving'.

De inhoud van dit werkboek vormde het uitgangspunt van een trainingsdag voor schoolbegeleiders die in samenwerking met het CPS werd gerealiseerd. Vooraf werd dit werkboek aan de schoolbegeleiders toegestuurd met het verzoek om in ieder geval de hoofdstukken over het voorgesprek, de observatie en het nagesprek te bestuderen. Centraal tijdens de trainingsdag stond het oefenen van de coachingsvaardigheden van het voor- en nagesprek en het schrijven van een beschrijvend verslag. De oefening in coachingsvaardigheden vond plaats via een rollenspel waarbij de schoolbegeleiders om beurten de rollen van coach, leraar en observator speelden. Na elke oefening werden leerpunten geformuleerd voor de eigen begeleidingspraktijk. Het maken van een beschrijvend verslag werd geoefend aan de hand van een video-opname van een les. Aan het eind van de trainingsdag werd het programma geëvalueerd, werden praktische suggesties gegeven en gaven de schoolbegeleiders enkele aanbevelingen voor de herziening van het werkboek. Het werkboek en het trainingspro-

gramma werd door de schoolbegeleiders positief beoordeeld. Bijna iedereen gaf aan te verwachten in de nabije toekomst gebruik te zullen maken van de geleerde coachingsvaardigheden.

6.5 Dataverzameling en -verwerking

Vóór de aanvang van het trainingsprogramma voerde de schoolbegeleider met een of twee leraren een coachingsgesprek uit. Deze voormeting viel in de periode mei-augustus 1993. Aan de schoolbegeleider werd verzocht een coachingsgesprek te voeren zoals men reeds eerder had gedaan of zoals men dacht dat het zou moeten. Half september vond de trainingsdag voor de experimentele groep plaats. Na afloop van de training voerde de schoolbegeleider met de betrokken leraren opnieuw een coachingsgesprek uit. Afsproken werd dat de gesprekken van de nameting afgerond zouden worden voor november. Het merendeel van de schoolbegeleiders kon deze afspraak niet gestand doen. Nadat deze schoolbegeleiders van zowel de experimentele groep als de controlegroep tweemaal telefonisch en eenmaal schriftelijk waren benaderd met het verzoek alsnog de coachingsgesprekken van de nameting uit voeren, werd de uiteindelijke sluitingsdatum vastgesteld op 15 februari 1994.

Van de schoolbegeleiders uit de experimentele groep voerde 36% een coachingsgesprek met twee leraren en 64% met één leraar. Van de schoolbegeleiders uit de controlegroep voerde 30% een coachingsgesprek uit met twee leraren en 70% met één leraar. Omdat de training was gericht op de uitvoering van een coachingscyclus werden van de getrainde schoolbegeleiders voor de nameting twee coachingsgesprekken per leraar verwacht, namelijk een voorgesprek en een nagesprek. Niet iedereen bleek de gehele coachingscyclus uitgevoerd te hebben. Van de experimentele groep voerde 78% de gehele cyclus uit. Uit de controlegroep voerde 20% van de schoolbegeleiders de gehele coachingscyclus uit. Bijna alle schoolbegeleiders voerden tijdens de voor- en nameting een nagesprek uit, dit wil zeggen dat zij bij hun coachingsgesprek gebruik maakten van observatiegegevens of van indrukken die zij in de klas hadden opgedaan.

Alle coachingsgesprekken werden door de schoolbegeleiders op cassetteband opgenomen

en door getrainde observatoren met behulp van de Schaal voor het Gebruik van Coachingsvaardigheden (SGCV) gecodeerd. Na elk coachingsgesprek kregen de gecoachte leraren van de schoolbegeleider in een gesloten envelop een vragenlijst waarin aan hen werd gevraagd hun mening te geven over het uitgevoerde coachingsgesprek. Deze vragen waren identiek aan die van de SGCV. Tevens kregen de leraren een aantal vragen voorgelegd over de mogelijke uitwerkingen van het coachingsgesprek op hun lesgedrag en hun functioneren. Deze vragenlijst voor de leraar werd rechtstreeks door de betrokken leraren naar het CPS opgestuurd, zodat de schoolbegeleider geen inzage had in de evaluaties van de gecoachte leraren.

Ten behoeve van de analyse werden voor die schoolbegeleiders die coachingsgesprekken uitvoerden met twee leraren de scores van de SGCV gemiddeld. Dit werd eveneens gedaan voor de scores van het voor- en nagesprek. Op deze wijze werd voor zowel de voormeting als de nameting voor elke schoolbegeleider één score berekend voor de toepassing van de coachingsvaardigheden. Bij de toetsing van verschillen tussen de scores van de voormeting en de nameting en tussen de experimentele groep en controlegroep werd een significantieniveau van 5% aangehouden (eenzijdig). De eenheid van analyse was de schoolbegeleider.

Voor een meer uitgebreide beschrijving van de opzet, het instrumentarium en de gevolgde werkwijze wordt verwezen naar De Jonghe en Van Wezel (1994).

7 Resultaten

Allereerst werd getoetst of de scores van de experimentele groep en de controlegroep op de voormeting significant van elkaar verschilden. Dit bleek niet het geval te zijn. Tussen de door de beoordelaars toegekende scores op de vier subschalen van de Schaal voor het Gebruik van Coachingsvaardigheden (SGCV) werden geen significante verschillen gevonden. Ook niet bij de vijf subschalen voor de door de leraren gegeven beoordelingen van de coachingsgesprekken. Dit betekende dat op de voormeting geen significante initiële verschillen tussen de experimentele en de controlegroep werden aangetroffen.

Tabel 1

Overzicht van de gemiddelde scores op de Schaal voor het Gebruik van Coachingsvaardigheden (SGCV) voor de experimentele groep en de controlegroep volgens het oordeel van de beoordelaars en de resultaten van de t-toetsen

Schaal voor het Gebruik van Coachingsvaardigheden	Exp. groep (n = 18)		Controle- groep (n = 10)		Verschilscore		t
	Voor	Na	Voor	Na	Exp. groep	Controle- groep	
SGCV-totaal	2.29	3.47**	2.15	2.14	1.18	-0.01	8.01**
Autonomie-bevordering	1.39	3.13**	1.32	1.32	1.75	-0.00	7.18**
Feedback:							
- voor- en nagesprek	2.99	2.52	2.78	2.58	-0.47	-0.20	-0.45
- nagesprek ¹	3.34	4.53**	3.23	3.03	1.19	-0.19	2.35*
Doelgerichtheid	2.57	4.12**	2.28	2.12	1.54	-0.16	7.88**
Verwerven van vertrouwen	3.60	4.52**	3.61	4.09	0.92	0.48	0.10

Opm.: ¹ Controlegroep N = 9

* $p < .05$; ** $p < .01$

Om te bepalen of de schoolbegeleiders uit de experimentele groep tijdens de nameting beter gebruik maakten van de coachingsvaardigheden uit het trainingsprogramma dan tijdens de voormeting, werd een t-toets voor gepaarde waarnemingen uitgevoerd. Verschillen tussen de experimentele groep en de controlegroep werden getoetst met behulp van een t-toets voor verschillen (scores op de nameting minus scores op de voormeting).

De gegevens betreffende de SGCV van de getrainde beoordelaars zijn samengevat in Tabel 1, alsmede de resultaten van de uitgevoerde t-toetsen.

Uit de gegevens van Tabel 1 blijkt dat de schoolbegeleiders uit de experimentele groep tijdens de nameting significant hoger scoorden op de SGCV en op de subschalen 'autonomie-bevordering', 'doelgerichtheid' en 'het verwerven van vertrouwen' dan op de voormeting. Dit betekent dat de schoolbegeleiders die het trainingsprogramma hadden gevolgd tijdens de nameting deze coachingsvaardigheden uit het coachingsmodel beter toepasten in hun coachingsgesprekken dan tijdens de voormeting. Alleen voor de subschaal 'feedback' (voor- en nagesprek) werd geen vooruitgang geconstateerd. Voor de schoolbegeleiders uit de controlegroep werden geen significante verschillen gevonden tussen de voor- en de nameting.

Bij vergelijking van de verschillen tussen de experimentele en de controlegroep door middel van verschillen bleek dat de schoolbegeleiders significant hoger scoorden op de SGCV als totaal dan de schoolbegeleiders uit de controlegroep. De uitkomsten op de subschalen

laten zien dat dit verschil vooral verklaard kon worden door de significante implementatie-effecten op de subschalen 'autonomie-bevordering' en 'doelgerichtheid'. Geen significant implementatie-effect werd gevonden voor 'feedback' (voor- en nagesprek) en 'het verwerven van vertrouwen'. Bij een vergelijking tussen de scores van de voor- en nameting van de experimentele groep werd een significante vooruitgang geconstateerd met betrekking tot het verwerven van vertrouwen. Bij een vergelijking tussen de scores van de experimentele groep en de controlegroep bleken hier geen significante verschillen te bestaan. Inspectie van de gegevens uit Tabel 1 laat zien dat ook de schoolbegeleiders uit de controlegroep vooruitgang boekten. Het tweede coachingsgesprek tijdens de nameting leidde waarschijnlijk tot een grotere vertrouwdheid tussen de schoolbegeleider en de gecoachte leraar. In de voormeting diende men nog aan elkaar te wennen.

Tijdens de nameting voerden de schoolbegeleiders twee soorten van coachingsgesprekken uit: een voorgesprek en een nagesprek. Het geven van feedback diende het meest tot zijn recht te komen in het nagesprek. Immers in het nagesprek kregen de leerkrachten informatie over het feitelijk gerealiseerd gedrag aan de hand van de verzamelde observatiegegevens in de klas. Om deze reden werd voor de subschaal 'feedback' een afzonderlijke analyse uitgevoerd op de nagesprekken van de nameting. Uit de analyse van de verschillen bleek dat de schoolbegeleiders uit de experimentele groep een significant beter gebruik maakten van de vaardigheden voor het geven van een goede

Tabel 2

Gegevens van enkele kenmerken van de schoolbegeleiders uit de experimentele groep en de resultaten van de t-toetsen op de verschillen met betrekking tot de Schaal voor het Gebruik van Coachingsvaardigheden (SGCV)

Controlevariabelen	SGCV-totaal		Autonomie bevordering		Verschilscores Feedback		Doelgerichtheid		Verwerven van vertrouwen	
	Gem.	t	Gem.	t	Gem.	t	Gem.	t	Gem.	t
Aantal jaren begeleidingservaring:										
weinig: 1-4 jaar (n = 10)	1.12	-0.81	1.60	-1.12	-0.42	0.14	1.42	-0.99	1.07	1.17
veel: 8-20 jaar (n = 8)	1.26		1.93		-0.53		1.70		0.73	
Kennis van het coachingsproces:										
ja (n = 11)	1.12	-0.90	1.68	-0.54	-0.65	-0.57	1.58	0.36	0.82	-0.82
nee (n = 7)	1.28		1.85		-0.18		1.48		1.07	
Enige ervaring met coaching:										
nee (n = 13)	1.19	0.10	1.70	-0.53	-0.37	0.38	1.50	-0.41	1.06	1.56
ja (n = 5)	1.17		1.88		-0.71		1.64		0.57	

feedback dan de schoolbegeleiders uit de controlegroep ($p < .05$).

Om te zien of het effect van het trainingsprogramma op de implementatie van de coachingsvaardigheden van de schoolbegeleiders uit de experimentele groep beïnvloed werd door enkele kenmerkende eigenschappen van deze schoolbegeleiders, werden enkele aanvullende analyses uitgevoerd. Onderzocht werd de invloed van de volgende controlevariabelen op de verschillen: het aantal jaren begeleidingservaring van de schoolbegeleider, de reeds eerder verworven kennis over coaching en de eerdere ervaring met coaching. De gegevens voor deze variabelen zijn samengevat in Tabel 2. Uit deze gegevens blijkt dat deze controlevariabelen geen effect hadden op de implementatie van de coachingsvaardigheden. Voor de onderscheiden groepen van schoolbegeleiders bleek het effect van het trainingspro-

gramma gelijk te zijn.

Tabel 3 geeft een overzicht van de scores die de gecoachte leraren gaven aan de uitgevoerde coachingsgesprekken van de schoolbegeleiders. De score voor de SGCV-totaal heeft betrekking op de gemiddelde score van de eerste vier subschalen. Deze vier subschalen werden ook gebruikt door de getrainde beoordelaars, zodat deze gegevens vergelijkbaar zijn. De vijfde subschaal was speciaal bedoeld voor de gecoachte leraren om hun inschatting te verkrijgen over de mogelijke effecten van het coachingsgesprek op het instructiegedrag.

Uit de gegevens van Tabel 3 blijkt dat volgens het oordeel van de gecoachte leraren de schoolbegeleiders uit de experimentele groep tijdens de nameting alleen significant hoger scoorden op de subschaal 'autonomie-bevordering' dan tijdens de voormeting. Geen significante verschillen werden gevonden tus-

Tabel 3

Overzicht van de gemiddelde scores op de Schaal voor het Gebruik van Coachingsvaardigheden (SGCV) voor de experimentele groep en de controlegroep volgens het oordeel van de gecoachte leraren en de resultaten van de t-toetsen

Schaal voor het Gebruik van Coachingsvaardigheden	Exp. groep (n = 18)		Controle-groep (n = 8)		Verschilscore		t
	Voor	Na	Voor	Na	Exp. groep	Controle-groep	
SGCV-totaal	3.60	3.69	3.67	3.77	0.08	0.09	-0.05
Autonomie-bevordering	3.11	3.53*	3.56	3.60	0.42	0.04	1.26
Feedback:							
- voor- en nagesprek	4.15	3.81	3.93	3.63	-0.34	0.30	-0.14
- nagesprek	4.31	3.81	3.89	3.75	-0.50	-0.13	-0.85
Doelgerichtheid	3.11	3.40	3.26	3.64	0.28	0.38	-0.32
Verwerven van vertrouwen	4.05	4.01	3.94	4.19	-0.04	0.25	-0.96
Waargenomen effecten op instructiegedrag	3.14	3.26	3.33	3.30	0.11	-0.04	0.58

* $p < .05$

sen de experimentele groep en de controle-groep. Opvallend is dat de leraren reeds tijdens de voormeting de uitvoering van de coachingsvaardigheden positief beoordeelden. Alle scores, ook die van de ingeschatte effecten van het coachingsgesprek op het eigen instructiegedrag, lagen boven het gemiddelde.

8 Discussie

De resultaten van het onderzoek geven aan dat het trainingsprogramma een positief effect heeft op het coachingsgedrag van de schoolbegeleiders. De gegevens van de getrainde beoordelaars laten zien dat de schoolbegeleiders die het trainingsprogramma hebben gevolgd hoger scoren op de Schaal voor het Gebruik van Coachingsvaardigheden als geheel dan schoolbegeleiders die niet aan het trainingsprogramma hebben deelgenomen. Meer in het bijzonder heeft het trainingsprogramma effect op de coachingsvaardigheden die te maken hebben met 'autonomie-bevordering' en 'doelgerichtheid'. 'Autonomie-bevordering' heeft betrekking op vaardigheden die de coach gebruikt om het vermogen van leraren te versterken om te reflecteren op het eigen instructiegedrag en om leraren te stimuleren zelf met alternatieven te komen om het instructiegedrag te verbeteren. 'Doelgerichtheid' heeft betrekking op het gericht zijn van de coach op de doelen van het coachingsgesprek: het formuleren van verbeteringsvoorstellen. De coach zorgt er voor dat het gesprek steeds terzake blijft. Met betrekking tot het toepassen van coachingsvaardigheden gericht op 'het verwerven van vertrouwen' boeken de getrainde schoolbegeleiders een significante vooruitgang, maar de schoolbegeleiders in de controlegroep gaan hier eveneens vooruit. Hierdoor is tussen beide groepen geen significant verschil gevonden. Deze vooruitgang kan op conto worden geschreven van een grotere vertrouwdheid tussen coach en leraar in de periode van de nameting dan in de periode van de voormeting. Het gesprek in de voormeting heeft de basis gelegd voor deze grotere vertrouwdheid.

Een algemeen effect voor de coachingsvaardigheden voor het geven van feedback is niet gevonden als de voor- en nagesprekken worden samengenomen. Wel is een significant

effect gevonden voor het geven van feedback tijdens het nagesprek. Dit is een positieve bevinding omdat het geven van goede feedback met name voor het nagesprek van belang is omdat de coach de resultaten meedeelt van de verkregen observaties. Aan de hand van het geven van goede feedback stimuleert de coach de leraar tot reflectie over zijn feitelijk gedrag in het licht van het gewenste gedrag. Goede feedback houdt in dat de coach in staat is constructieve gegevens uit de lesobservatie naar voren te brengen, dat hij zo objectief mogelijk probeert te beschrijven wat hij heeft waargenomen, en dat hij het aantal aspecten van het lesgedrag dat voor verbetering in aanmerking komt weet te beperken tot een hanteerbaar geheel. In het nagesprek blijken de schoolbegeleiders die aan het training hebben deelgenomen deze kenmerken van goede feedback beter te realiseren dan schoolbegeleiders die niet aan de training hebben deelgenomen.

Hoewel alle mogelijke moeite is gedaan om een zo groot mogelijke experimentele groep samen te stellen in relatie tot de doelen van een echte vaardigheidstraining, is het niet gelukt om alle beoogde 25 schoolbegeleiders bij het onderzoek te betrekken. Door drukke werkzaamheden meldden zich voor de trainingsdag nog enige schoolbegeleiders af en twee schoolbegeleiders zagen geen kans een coachingsgesprek te voeren. Dit reduceerde het aantal schoolbegeleiders in de experimentele groep tot 18. Ook twee schoolbegeleiders die zich aangemeld hadden voor de tweede trainingsgroep en daarmee voor de controlegroep zagen in de periode van de nameting geen kans een coachingsgesprek te voeren. Hierdoor bestond de uiteindelijke controlegroep uit 10 schoolbegeleiders.

Het heeft veel moeite en overredingskracht gekost om een aantal schoolbegeleiders uit de experimentele groep tijdens de nameting de gehele coachingscyclus te laten uitvoeren. Uitvoering ervan hield in een voorgesprek, een observatie in de klas en een nagesprek. Uit het rappelleren is naar voren gekomen dat voor veel schoolbegeleiders het begeleiden van leraren in de klas nog geen gangbare praktijk was. Het kostte deze schoolbegeleiders veel moeite om tijdens de periode van de voor- en nameting leraren te vinden die bereid waren zich te laten coachen, omdat het ook voor de leraren een

nieuwe ervaring was.

In tegenstelling tot de getrainde beoordeelaars, constateerden de gecoachte leraren geen verschillen tussen getrainde en niet-getrainde schoolbegeleiders met betrekking tot de uitvoering van coaching. Voor dit verschil in beoordeling kunnen mogelijk de volgende verklaringen dienen. De getrainde observatoren kenden de inhoud van het werkboek en de daarop gegeven omschrijvingen van de coachingsvaardigheden door en door. Bovendien hadden zij door de training in het hanteren van de SGCV een nauwkeurig beeld van wat de betreffende coachingsvaardigheid inhield. Voor de gecoachte leraren waren de coachingsvaardigheden veel minder specifiek. Zij vulden de SGCV in op grond van de eenregelige omschrijvingen van de coachingsvaardigheden. Deze invulling op grond van meer globale impressies kan er toe geleid hebben dat verschillen in de uitvoering van de coachingsvaardigheden gemaskeerd werden. Verder kan het optreden van een 'Hawthorne-effect' niet uitgesloten worden geacht. De betrokken leraren waren zo vereerd met de extra aandacht van de schoolbegeleiders voor de praktijk van hun lesgeven, dat zij deze aandacht in hun scores tot uitdrukking hebben gebracht. Dit blijkt uit de relatief hoge scores op de voormeting. Als gevolg hiervan is mogelijk een plafond-effect opgetreden: het was bijna niet meer mogelijk de relatief hoge scores van de voormeting in de nameting te overtreffen. Verder dient bedacht te worden dat de schoolbegeleiders zelf hun leraren voor het coachingsproces hebben uitgekozen. Deze leraren hadden een positieve relatie met hun schoolbegeleider. Een selectie-effect kan niet uitgesloten worden geacht. Indien de schoolbegeleiders niet zelf hun te coachen leraren hadden gekozen, waren mogelijk wel verschillen tussen getrainde en niet-getrainde schoolbegeleiders geconstateerd.

Wel blijkt uit de oordelen van de leraren dat zij het effect van coaching op de verbetering van hun instructiegedrag redelijk hoog inschaten. Uit de gegevens van de voormeting blijkt dat het nadenken over het eigen instructiegedrag met een coach en het in-dialoog-formuleren van intenties voor mogelijke verbeteringen de kansen op daadwerkelijke veranderingen vergroten. Dit betekent een krachtige ondersteuning van het belang van coaching.

In dit onderzoek is slechts het implementatie-effect van het trainingsprogramma onderzocht. Schoolbegeleiders bleken in staat een aantal belangrijk geachte coachingsvaardigheden in de praktijk toe te passen. Of deze coachingsvaardigheden en het gekozen coachingsmodel leiden tot veranderingen in het denken en handelen van leraren dient in toekomstig onderzoek nader onderzocht te worden.

Literatuur

- Association for Supervision and Curriculum Development (ASCD) (1988). *Another set of eyes: Conferencing skills*. (Trainers manual). Alexandria: Author.
- Baker, R. G. (1983). *The contribution of coaching to transfer of learning: An extension study*. Unpublished doctoral dissertation, University of Oregon.
- Bennett, B. (1987). *The effectiveness of staff development training practices: A meta-analysis*. Unpublished doctoral dissertation, University of Oregon.
- Commissie Toekomst Leraarschap. (1993). *Een beroep met perspectief: De toekomst van het leraarschap*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.
- Cogan, M. L. (1973). *Clinical Supervision*. Boston: Houghton Mifflin.
- Costa, A. L., & Garmston, R. J. (1994). *Cognitive coaching: A foundation for Renaissance schools*. Norwood: Christopher-Gordon.
- Goldhammer, R. (1969). *Clinical supervision: Special methods for the supervision of teachers*. New York: Holt, Rinehart and Winston.
- Goldhammer, R., Anderson, R. H., & Krajewski, R. J. (1980). *Clinical supervision: Special methods for the supervision of teachers* (2nd edition). New York: Holt, Rinehart and Winston.
- Houtveen, A. A. M. (1990). *Begeleiden van vernieuwingen*. (Academisch Proefschrift, Rijksuniversiteit Utrecht) De Lier: Academisch Boeken Centrum.
- Houtveen, A. A. M., Booij, N., & Jong, R. de (1995a). *Effectieve begeleiding en leerkrachtgedrag*. Utrecht: ISOR Onderwijsonderzoek.
- Houtveen, A. A. M., Booij, N., & Jong, R. de (1995b). *Resultaten van het evaluatie-onderzoek naar het Landelijk Project Schoolverbetering*. Utrecht: ISOR Onderwijsonderzoek.

- Hunter, L. (1986). Let's eliminate the preobservation conference. *Educational Leadership*, 43(6), 69-70.
- Hunter, M., & Russel, D. (1990). *Mastering coaching and supervision*. El Segundo: TIP Publications.
- Jonghe, H. de, & Wezel, E. van (1994). *Coaching: Een onderzoek naar de effecten van het trainingsprogramma 'De schoolbegeleider als coach'*. (Doctoraal scriptie) Nijmegen: Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen.
- Joyce, B., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership*, 37(6), 379-385.
- Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational Leadership*, 40(10), 4-8.
- Joyce, B., & Showers, B. (1988). *Student achievement through staff development*. New York: Longman.
- Kinlaw, D. C. (1989). *Coaching for commitment: Managerial strategies for obtaining superior performance*. San Diego: University Associates.
- Nolan, J., Hawkes, B., & Francis, P. (1993). Case studies: Windows onto clinical supervision. *Educational Leadership*, 51(2), 52-56.
- Pajak, E. (1993). *Approaches to clinical supervision: Alternatives for improving instruction*. Norwood: Christopher-Gordon.
- Pavan, B. N. (1985, April). *Clinical supervision: Research in schools utilizing comparative measures*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago. (ERIC Document Reproduction Service No. ED 255 516)
- Pavan, B. N. (1991, April). *Determining the usages of clinical supervision*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago. (ERIC Document Reproduction Service No. ED 336 374)
- Pavan, B. N. (1993). Examining clinical supervision practice. In R. H. Anderson & K. J. Snyder (Eds.), *Clinical supervision: Coaching for higher performance* (pp. 135-154). Lancaster: Technomic.
- Roelofs, E. (1993). *Teamgerichte nascholing en coaching: Een experimentele studie in scholen met combinatieklassen*. (Academisch proefschrift, Katholieke Universiteit Nijmegen) Nijmegen: Universiteitsdrukkerij.
- Roelofs, E., Veenman, S., & Raemaekers, J. (1993). Teamgerichte nascholing en coaching: Effecten op het didactisch handelen van leerkrachten in combinatieklassen. *Pedagogische Studiën*, 70, 352-367.
- Roelofs, E., Veenman, S., & Voeten, M. (1994). Effecten van een teamgerichte nascholing en van coaching op leerprestaties. *Pedagogische Studiën*, 71, 442-456.
- Schön, D. A. (1983). *The reflective practitioner*. New York: Basic Books.
- Sergiovanni, T. J., & Starratt, R. J. (1983). *Supervision: Human perspectives*. New York: McGrawHill.
- Showers, B. (1982). *Transfer of training: The contribution of coaching*. Eugene: University of Oregon, Center for Educational Policy and Management. (ERIC Document Reproduction Service No. ED 231 035)
- Snippe, J. (1991). *In-service training voor leerkrachten*. (Academisch proefschrift, Rijksuniversiteit Groningen) Groningen: RION.
- Sparks, G. M., & Bruder, S. (1987, April). *How school-based peer coaching improves collegiality and experimentation*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington.
- Veenman, S., Jonghe, H. de, & Wezel, E. van. (1995). *Het coachen van leerkrachten: Een werkboek voor het begeleiden van leerkrachten*. Hoevelaken: Christelijk Pedagogisch Studiecentrum.
- Veenman, S., Lem, P., Roelofs, E., & Nijssen, F. (1994). *Effectieve instructie en doelmatig klasmanagement: Een schoolverbeteringsprogramma voor enkelvoudige en combinatieklassen* (5de druk). Amsterdam/Lisse: Swets & Zeitlinger.
- Verkerk, J. (1987). *Klasseconsultatie: Teambegeleiding als klasse-werk*. Hoevelaken: Christelijk Pedagogisch Studiecentrum.
- Westerop, H. (1993). *Kijk op eigen kunnen: Collegiale consultatie*. Hoevelaken: Christelijk Pedagogisch Centrum.

Manuscript aanvaard 29-8-1995

Auteurs

S. Veenman is als universitair hoofddocent verbonden aan de vakgroep Onderwijskunde van de Katholieke Universiteit te Nijmegen

H. de Jonghe is als onderwijskundige verbonden aan de Onderwijsbegeleidingsdienst Iselinge te Doetichem.

E. van Wezel is als onderwijskundige verbonden aan de Nederlandse School in Nairobi, Kenia.

Correspondentie-adres: Vakgroep Onderwijskunde, K.U. Nijmegen, Postbus 9104, 6500 HE Nijmegen.

Abstract

Evaluation of a training program for coaching teachers

S. Veenman, H. de Jonghe & E. van Wezel. *Pedagogische Studiën*, 1996, 73, 170-183.

This study describes the implementation effects of a program for the training of coaching skills with Dutch school counsellors. Coaching is a form of in-class support intended to provide teachers with feedback on their own functioning and thereby stimulate self-reflection and self-analysis in order to improve instructional effectiveness. To assist school counsellors in their attempts to coach teachers, a training program was designed: 'The school counsellor as coach'. A quasi-experimental, treatment-control group investigation was set up to test the effects of this program. Based on the pre- and post-training ratings of coaching conferences, a significant treatment effect was found for the coaching skills concerned with the development of autonomy (empowerment), feedback, and purposefulness. The pre- and post-training ratings from coached teachers showed no significant treatment effects. The positive ratings from the coached teachers showed them to perceive the coaching conferences as improving their instruction.

Effecten van voorschoolse stimuleringsprogramma's: een review van reviews

H. Blok en P. P. M. Leseman

Samenvatting

Bij de bestrijding van onderwijsachterstanden worden voorschoolse stimuleringsprogramma's in preventieve zin ingezet. In dit artikel wordt de vraag behandeld hoe sterk het effect van zulke programma's op de ontwikkeling van kinderen is. Ter beantwoording zijn tien reviews – verschenen tussen 1983 en 1993 – geanalyseerd. De reviews zijn op vier criteria beoordeeld: de selectie van de bronstudies, de codering van de kenmerken van de studies, de manier waarop de resultaten van de studies zijn geïntegreerd en de correspondentie tussen de gepresenteerde data en de getrokken conclusies. Acht van de tien reviews blijken ten aanzien van deze criteria ernstige tekorten te vertonen. De resterende twee reviews laten de volgende conclusies toe: a. na afloop van de interventie laten IQ-testen en schoolvorderingen-toetsen een gemiddelde winst van ongeveer een halve sigma zien; b. de winst op het sociale domein bedraagt ongeveer een kwart sigma; c. in de meeste gevallen is de winst na één tot twee jaar niet meer aantoonbaar; d. effecten op de schoolloopbaan worden in slechts enkele studies aangetoond; e. er is weinig bekend over de consequenties van vormgevingskenmerken voor de effectiviteit van de programma's. Ten slotte, aangezien het meeste onderzoek tussen 1960 en 1980 is uitgevoerd bij Amerikaanse achterstandskinderen, is de generaliseerbaarheid van deze conclusies naar de huidige Nederlandse situatie waarschijnlijk beperkt.

1 Inleiding

De aandacht voor voorschoolse educatie is de afgelopen jaren duidelijk gegroeid. Eén van de factoren hierbij is het toenemend gebruik van voorzieningen als kinderdagverblijven en peuterspeelzalen. Een andere factor is de teleur-

stelling over de resultaten van de strijd tegen onderwijsachterstanden. Steeds meer wint het inzicht veld dat onderwijsachterstanden beter voorkomen dan verholpen kunnen worden. De veronderstelling is dat voorschools onderwijs ervoor kan zorgen dat kinderen die op hun vierde jaar de basisschool betreden, een min of meer gelijke start kunnen maken. Met name kinderen die thuis in een achterstandssituatie opgroeien, zouden van voorschools onderwijs profiteren.

Naar de effectiviteit van voorschoolse stimuleringsprogramma's is – in het buitenland – inmiddels heel wat onderzoek verricht. Gerekend vanaf de jaren zestig loopt het aantal afzonderlijke studies in de honderden. Het is voor individuele onderzoekers feitelijk onmogelijk om zo'n grote hoeveelheid onderzoek te overzien. Gelukkig zijn er overzichtsstudies verschenen, die een goed startpunt kunnen vormen om tot een verantwoorde afweging te komen van de effecten van voorschools onderwijs.

Wij hebben de opdracht gekregen alle effectonderzoek, ná 1985 verschenen, te reviewen¹. Dat gebeurt in twee stappen. In de eerste, voorbereidende stap – waarover we hier rapporteren – hebben we de reeds beschikbare, recente reviews geanalyseerd. In de tweede stap, waarover we in de nabije toekomst hopen te rapporteren, komen we aan onze eigenlijke opdracht toe, namelijk het reviewen van het na 1985 gepubliceerde effectonderzoek.

Wij hebben in totaal tien reviews – verschenen tussen 1983 en 1993 – geanalyseerd. Onze analyse is alleen gebaseerd op de reviews zelf. We zijn dus niet teruggegaan naar de bronstudies. In eerste instantie ging onze belangstelling vooral uit naar de conclusies die de auteurs aan hun reviews verbonden. Naarmate wij meer reviews verzamelden en steeds vaker constateerden dat de daarin vermelde conclusies op belangrijke punten divergeerden, kregen we meer belangstelling voor de kwaliteit

van de reviews. Dit artikel is dus meer dan alleen een beschrijvende samenvatting van eerdere reviews. Het is in belangrijke mate ook een kritische analyse van de wijze waarop de reviewers zich van hun taak hebben gekwet.

2 Probleemstelling en opzet

2.1 Probleemstelling

In deze bijdrage onderscheiden we een primaire en een secundaire vraag. De primaire vraag is die naar de effecten van voorschoolse stimuleringsprogramma's. De secundaire vraag betreft de kwaliteit van de sinds 1980 verschenen reviews over dit onderwerp. Beide vragen worden toegelicht.

De *primaire* vraag luidt: 'Hoe sterk is het effect van voorschoolse stimuleringsprogramma's op de ontwikkeling van kinderen?' Bij de uitwerking van deze vraag hebben we de volgende specificaties aangebracht. Met de term stimuleringsprogramma's wordt op programma's gedoeld die expliciet tot doel hebben de ontwikkeling van kinderen te stimuleren. Het gaat hier om programma's die een uitgesproken educatieve doelstelling hebben. Zulke stimuleringsprogramma's onderscheiden zich van programma's die als hoofddoel hebben, kinderen een opvang te bieden voor de momenten dat ouders geen zorg willen of kunnen bieden. Zulke opvangprogramma's laten we buiten beschouwing, evenals programma's die alleen op ouders gericht zijn, bijvoorbeeld ter ondersteuning van hun opvoedingsgedrag.

Het merendeel van de stimuleringsprogramma's is 'center-based', hetgeen impliceert dat de kinderen het programma buiten hun eigen huis volgen. In sommige gevallen bevatten programma's naast een 'center-based' component ook een 'home-based' component, waarin bijvoorbeeld zaken als thuisbezoeken plaatsvinden.

De kwalificatie 'voorschoolse' heeft binnen verschillende schoolsystemen een verschillende invulling. Maar aangezien de gereviewde literatuur vrijwel uitsluitend de Verenigde Staten betreft, gaat het om kinderen in de leeftijd vanaf nul tot en met vier jaar. Vanaf vijf jaar bezoeken de meeste Amerikaanse kinderen een 'Kindergarten', een onderwijssoort die in enkele opzichten vergelijkbaar is met

Nederlandse kleuterklassen.

Aan de ontwikkeling van kinderen zijn vanzelfsprekend heel diverse facetten te onderscheiden. In de onderwijskundige literatuur wordt veelal een driedeling gemaakt: het cognitieve domein (waaronder intelligentie, ontlukende geletterdheid en ontlukende gecijferdheid), het sociaal-emotionele domein en het motorische domein (Seifert & Hoffnung, 1994). Wij hebben binnen alle drie domeinen naar relevante reviews gezocht, maar de aandacht is, zoals bekend, ongelijk verdeeld. De meeste aandacht is in het verleden uitgegaan naar het cognitieve domein, op grote afstand gevolgd door de beide andere domeinen.

De laatste specificatie betreft de kinderen voor wie de programma's bedoeld zijn. In eerste instantie wordt deze leeftijdsgroep als ongedeelde populatie beschouwd. Maar binnen deze populatie is een subpopulatie die van onderwijskundigen bijzondere aandacht gekregen heeft: de kinderen die, in Amerikaanse termen gesproken, om economische, culturele of sociale redenen in een risicosituatie verkeren. Deze subpopulatie komt in belangrijke opzichten overeen met de doelgroep van het Nederlandse onderwijsvoorrangsbeleid. In feite zijn veel van de ontwikkelde stimuleringsprogramma's op deze subpopulatie gericht. Uitgesloten hebben wij een andere subpopulatie: kinderen die om medische of paramedische redenen een risicogroep vormen, waaronder doven, blinden/slechtzienden, te vroeg geboren kinderen, mentaal geretardeerden, kinderen met spraak/taalstoornissen, kinderen van drugsverslaafden, enzovoorts.

Daarmee komen we toe aan de *secundaire* vraag. In de meeste leerboeken wordt onderzoekers aangeraden de wetenschappelijke kwaliteit van hun bronnen grondig te analyseren, voordat men conclusies hieruit overneemt. Die aanbeveling geldt niet alleen voor primaire bronnen zoals onderzoeksrapporten en artikelen over afzonderlijke studies, maar ook voor secundaire bronnen zoals reviews. Voor secundaire bronnen is dit evenwel vaak moeilijk te doen. Het ontbreekt menig onderzoeker aan tijd om terug te gaan naar de oorspronkelijke bronnen, en, het vervolgens combineren van de gegevens uit soms tientallen onderzoeken is een heel bewerkelijke zaak. Omdat er in veel gevallen bovendien maar één review beschik-

baar is, heeft een individuele onderzoeker meestal geen andere keus dan zich geheel en al te verlaten op de conclusies die de reviewer aan het einde van zijn studie getrokken heeft. In de praktijk ziet men dan ook vaak dat conclusies uit reviews letterlijk en zonder kritische kanttekeningen worden overgenomen.

Dit is een risicovolle werkwijze. In de enkele gevallen dat reviews kritisch onder de loep zijn genomen, zijn de uitkomsten desastreus. Zo laat Slavin (1984) zich uiterst kritisch uit over een achttal belangrijke meta-analyses uit de eerste helft van de jaren tachtig: "... every one I examined in detail contained sources of bias serious enough, in my judgment, to call into question one or more major conclusions" (p. 8). Een ander voorbeeld vormen White, Taylor en Moss (1992) die de vloer aanvegen met zes reviews naar het effect van 'parent involvement' in voorschoolse stimuleringsprogramma's. Ten aanzien van de veronderstelling dat het betrekken van ouders de effectiviteit van de programma's gunstig beïnvloedt, concluderen zij: "The data cited by prominent reviewers of the early intervention research literature are largely contradictory or irrelevant to such a hypothesis" (p. 118).

In ons geval verkeerden we in de gelukkige omstandigheid dat er over één terrein verscheidene min of meer recente reviews beschikbaar waren. Dat biedt de gelegenheid ze naar hun aanpak en resultaten te vergelijken. De secundaire onderzoeksvraag luidde: 'Hoe degelijk zijn de reviews - ten aanzien van het in de primaire vraagstelling aangeduide probleemgebied - uitgevoerd?' Bij de beantwoording van deze vraag is het onmogelijk gebleken terug te gaan naar de oorspronkelijke of primaire studies. De werklast die hiermee gepaard zou gaan, lag ver boven onze mogelijkheden. Bij de beoordeling van de reviews hebben we alleen gebruik gemaakt van gegevens die in de reviews zelf vermeld zijn. De wijze waarop we te werk gegaan zijn, wordt in paragraaf 2.3 nader uitgewerkt.

2.2 Selectie van reviews

Wij hebben een poging gedaan alle relevante reviews van na 1980 bij onze review te betrekken. Reviews van vóór het aangegeven jaar zijn in zekere zin te beschouwen als behorende tot een eerdere generatie. Hun conclusies zijn

overgenomen, en, indien nodig, gecorrigeerd in latere overzichten. Dit geldt bijvoorbeeld voor de eerste review van Head Start projecten (Westinghouse Learning Corporation, 1969), die inmiddels is achterhaald door de review van McKey, Condelli, Ganson, Barrett, McConkey en Plantz (1985).

Om de reviews op het spoor te komen, zijn searches uitgevoerd in de bestanden van ERIC en PsycLit. Tevens zijn jaargangen van de relevante tijdschriften van de 'American Educational Research Association' doorgenomen, evenals handboeken en jaarboeken, zoals de 'Annual Review of Psychology' en de 'Review of Research in Education'. Onze speurtocht is beperkt gebleven tot de Engelstalige literatuur. Reviews die als prepublicatie zijn uitgebracht (bijvoorbeeld Harrell, 1983), zijn buiten beschouwing gebleven. Evenmin is Lazar en Darlington (1982) opgenomen, omdat het 'Consortium for Longitudinal Studies' (1983) nieuwe data bevat - namelijk op basis van een extra follow-up in 1980 - die in Lazar en Darlington (1982) niet beschreven worden.

Ten aanzien van de reviews met een bredere vraagstelling dan wij hier hanteren, is als volgt gehandeld. Indien in zulke reviews het voor ons relevante gedeelte onderscheidbaar was van het gedeelte dat onze vraagstelling te buiten ging, hebben wij ze toch opgenomen. Dit geldt bijvoorbeeld voor White (1985-1986) en Faran (1990) die zowel programma's voor achterstandskinderen als voor gehandicapte kinderen bespreken. De niet van toepassing zijnde delen zijn buiten beschouwing gebleven. Evenwel, in de review van Clark (1988), waarin alle Engelse onderzoek naar voorschoolse stimulering opgenomen is, bleek het onmogelijk een gedeelte af te zonderen dat betrekking had op de effectiviteit van programma's. Deze review is daarom door ons niet opgenomen.

Uiteindelijk zijn tien reviews uitgekozen. In de literatuurlijst zijn deze met een asterisk gemarkeerd. Naar ons beste weten hebben we daarmee alle reviews die aan de gestelde criteria voldeden bij onze analyse betrokken.

2.3 Verdere werkwijze

Bij de beoordeling van de reviews zijn vier hoofdcriteria onderscheiden: selectie, codering, integratie en conclusies. Deze criteria zijn ontleend aan door Kulik en Kulik (1989) op-

gestelde richtlijnen. Per hoofdcriterium wordt een korte toelichting gegeven.

Selectie

Een goede review voldoet aan een aantal eisen wat betreft de selectie van de studies die als primaire bronnen worden gebruikt. Hoofdpunt is dat de reviewer expliciete criteria hanteert om studies wel of niet op te nemen. Deze criteria moeten afgeleid zijn uit de centrale vraag van de review. De criteria zijn als het ware de meest concrete operationalisatie van de onderzoeksvraag, waarin bijvoorbeeld de afhankelijke en onafhankelijke variabelen nader worden gespecificeerd. Aanvullend kunnen criteria geformuleerd worden over de methodologische kwaliteit van de studies, maar het nut hiervan is omstreken. Vanzelfsprekend behoren de gehanteerde criteria vermeld te worden en behoort de reviewer er zich aan te houden. Binnen deze criteria dient naar volledigheid gestreefd te worden.

Codering

In een goede review worden relevante kenmerken van de studies gecodeerd. Daardoor is het mogelijk de relaties tussen kenmerken van de studies en de uitkomsten statistisch te analyseren. Relevante kenmerken betreffen de aard van de afhankelijke variabelen en de gehanteerde operationalisaties, de aard en de lengte van de experimentele behandeling, de resultaten van de studie en een aantal aspecten van de experimentele opzet (bijvoorbeeld het type opzet). Daarnaast kunnen variabelen die de publicatiestatus betreffen (bijvoorbeeld het jaar van publicatie, het type publicatie), worden gecodeerd. Het spreekt voor zich dat de reviewer aandacht moet besteden aan de betrouwbaarheid van de coderingsprocedures.

Integratie

Bij de integratie van de studie-uitkomsten zullen veelal één of meer maten voor de sterkte van de gevonden relaties een rol spelen. Dit kan een correlatie-coëfficiënt zijn, of een op één of andere wijze gestandaardiseerde maat voor de effectgrootte. Bij het combineren van de gekozen maten moeten de juiste correcties worden toegepast. Voorts moet verhinderd worden dat studies met veel afhankelijke variabelen een te groot gewicht krijgen. Eenheid van analyse

dient de studie te zijn, niet de effectgrootte. In voorkomende gevallen dienen daarom in een voorfase de uitkomsten per studie 'gepooled' te worden.

Conclusies

Een laatste eis is dat reviewers geen conclusies trekken die niet door hun analyses worden ondersteund. De conclusies dienen door de beschreven uitkomsten ondersteund te worden.

3 Resultaten

3.1 De eerste oriëntatie

Tabel 1 biedt een eerste oriëntatie op de tien reviews die onder de loep genomen zijn. Wat betreft het aantal primaire studies waarop de reviews zich baseren, zijn er grote verschillen. Stallings en Stipek (1986) bespreken - naast de review van het Consortium uit 1983 - slechts twee primaire studies. Het andere uiterste is McKey et al. (1985) die 210 studies bespreken. Dat lijkt een gigantisch aantal. Maar van deze 210 studies bleken er slechts 76 te zijn die voldoende gegevens bevatten voor een meta-analyse.

We hebben een poging gedaan de overlap tussen de reviews wat betreft de primaire bronnen te bepalen. Maar omdat niet alle reviews een volledige lijst van hun bronnen bevatten, is dit maar gedeeltelijk gelukt. Met name de reviews van White (1985-1986) en van White, Taylor en Moss (1992) missen een lijst met gereviewde bronnen. De review van het 'Consortium for Longitudinal Studies' (1983) is in zekere zin als een belangrijke kern te beschouwen. In alle latere reviews, waarschijnlijk ook in die van White (1985-1986) en van White, Taylor en Moss (1992), komen één of meer van de door het 'Consortium for Longitudinal Studies' gereviewde bronnen terug. Bovendien zijn er twee reviews die een bespreking wijden aan de review van het 'Consortium for Longitudinal Studies' zelf, namelijk Stallings en Stipek (1986) en Sylva en Wiltshire (1993). De review van McKey et al. (1985) blijkt tamelijk los te staan van de latere reviews. Voor zover wij konden nagaan, is van de hierin opgenomen bronstudies geen enkele in latere reviews terecht gekomen. De review zelf wordt alleen in Karweit (1989) aan de orde ge-

Tabel 1*Beschrijvende kenmerken van tien reviews*

Reviews	Aantal studies	Overlap met eerdere reviews ^b	Leerlingenpopulatie	Interventieperiode
'Consortium for Longitudinal Studies' (1983)	11	niet van toepassing	Amerikaans, achterstands-groeperingen, 0-5 jr.	tussen 1961 en 1969
Clarke-Stewart & Fein (1983)	ca. 80 ^{a,c}	8 maakten deel uit van 'Consortium for Longitudinal Studies' (1983)	Niet vermeld, vermoedelijk voornamelijk Amerikaans, gemengd qua SES status, 0-5 jr.	tussen 1955 en 1980
McKey et al. (1985)	210	2 studies maakten deel uit van 'Consortium for Longitudinal Studies' (1983)	Amerikaans, achterstands-groeperingen, 3-4 jr.	tussen 1965 en 1980
White (1985-1986)	99 ^a	onbekend (geen lit. lijst beschikbaar)	Amerikaans, achterstands-groeperingen, 7-5 jr.	merendeels tussen 1960 en 1982
Stallings & Stipek (1986)	3 (w.o. één review) ^a	volledige overlap met 'Consortium for Longitudinal Studies' (1983)	Amerikaans, achterstands-groeperingen, 0-5 jr.	tussen 1961 en 1964
Goldring & Presbey (1986)	8	volledige overlap met 'Consortium for Longitudinal Studies' (1983)	Amerikaans, achterstands-groeperingen, 0-5 jr.	tussen 1961 en 1969
Karweit (1989)	6 (w.o. één review) ^a	de review is McKey et al.; vier van de vijf studies maakten deel uit van 'Consortium for Longitudinal Studies' (1983)	Amerikaans, achterstands-groeperingen, 3-4 jr.	tussen 1961 en 1980
Farran (1990)	24 ^a	7 studies maakten deel uit van 'Consortium for Longitudinal Studies' (1983)	voornamelijk Amerikaanse achterstandsgroeperingen, 0-5 jr.	tussen 1961 en 1984
White, Taylor & Moss (1992)	97 ^a	onbekend (geen lit. lijst beschikbaar)	Amerikaans, achterstands-groeperingen, 7-5 jr.	tussen 1961 en 1982
Sylva & Wiltshire (1993)	17 (w.o. drie reviews)	1 studie maakte deel uit van 'Consortium for Longitudinal Studies' (1983)	deels Amerikaans, deels Engels, deels achterstands-groeperingen, deels dwarsdoorsnede uit bevolking	tussen 1961 en 1990

^a Vermeld wordt het aantal studies binnen het voor ons relevante deel.^b Bij dit kenmerk wordt de review van het 'Consortium for Longitudinal Studies' (1983) als de eerste beschouwd.^c Het aantal studies is niet precies te bepalen.

steld. Ten slotte, alleen Farran (1990) en Sylva en Wiltshire (1993) bevatten een beduidend aantal nieuwe - niet in het 'Consortium for Longitudinal Studies' (1983) of in McKey et al. (1985) besproken - bronnen.

Ook al zijn de reviews zelf van redelijk recente datum, de voorschoolse stimuleringsprogramma's waarover gerapporteerd wordt zijn beduidend ouder. De programma's die besproken worden in de review van het 'Consortium

for Longitudinal Studies' (1983) zijn uit de jaren zestig, die van McKey et al. (1985) uit 1965 tot aan 1980. Alleen Sylva en Wiltshire (1993) bespreken programma's die minder dan tien jaar geleden zijn uitgevoerd.

Kenmerkend is het vrijwel exclusieve accent op programma's die in de Verenigde Staten zijn ontwikkeld voor kinderen in achterstandssituaties. Gezien de heersende Amerikaanse opvattingen gaat het om kinderen uit

Tabel 2

Beoordelingen van tien reviews (+ = positief oordeel; - = negatief oordeel of geen oordeel mogelijk)

Reviews	Criteria			
	Selectie	Codering	Integratie	Conclusies
'Consortium for Longitudinal Studies' (1983)	+	+	+	-
Clarke-Stewart & Fein (1983)	-	-	-	-
McKey et al. (1985)	+	+	+	+
White (1985-1986)	-	-	-	+
Stallings & Stipek (1986)	-	-	-	-
Goldring & Presbey (1986)	+	-	-	-
Karweit (1989)	-	-	-	-
Farran (1990)	-	-	-	-
White, Taylor & Moss (1992)	-	-	-	+
Sylva & Wiltshire (1993)	-	-	-	-

arme gezinnen, voornamelijk als 'black' aangeduid. Clarke-Stewart en Fein (1983) en Sylva en Wiltshire (1993) zijn de enige twee reviews die niet exclusief focussen op programma's voor kinderen in achterstandssituaties. Sylva en Wiltshire bespreken enkele programma's voor een dwarsdoorsnede uit de Engelse bevolking, maar maken in hun review niet duidelijk of de aard van de populatie waar de interventieprogramma's voor bedoeld zijn, van invloed is op de effectiviteit ervan.

3.2 De reviews afzonderlijk besproken

In deze paragraaf worden de afzonderlijke reviews kort besproken. Een meer uitgebreide bespreking is nu eenmaal niet mogelijk binnen de ruimte die een tijdschrift kan bieden. In de besprekingen wordt verwezen naar beoordelingen ten aanzien van de vier in paragraaf 2.3 onderscheiden criteria voor kwaliteit. De oordelen zijn in Tabel 2 samengevat.

'Consortium for Longitudinal Studies' (1983). Halverwege de jaren zeventig werd een consortium gevormd door onderzoekers die de lange-termijneffecten van hun vroege-interventieprogramma's - uitgevoerd in de jaren zestig - wilden vaststellen. In het consortium werden elf (van de in totaal 15 geïdentificeerde) afzonderlijke studies ingebracht. Overeengekomen werd zowel in 1976 als in 1980 follow-up data te verzamelen. Ondanks alle in het werk gestelde pogingen zoveel mogelijk van de oorspronkelijke proefpersonen te bereiken, bedroeg het aantal proefpersonen in 1980 nog slechts ongeveer 1100, dat wil zeggen minder dan één derde van het aantal oorspronkelijke proefpersonen. Bij de laatste follow-up varieerden de proefpersonen qua leeftijd van

10 tot meer dan 20 jaar. Ook inhoudelijk varieerden de programma's aanzienlijk (onder andere wat betreft de leeftijd waarop de interventie destijds plaats vond, de betrokkenheid van de ouders en de setting). In tegenstelling tot de andere hier te bespreken reviews zijn in deze review de oorspronkelijke data geanalyseerd. Het mediane effect, onmiddellijk na de interventie, werd geschat op 7 IQ-punten (gebaseerd op 7 van de 11 projecten), afnemend tot 3 IQ-punten drie tot vier jaar later. Op schoolvorderingen werden nauwelijks verschillen aangetoond. Opvallend is dat er op schoolloopbaanvariabelen wel verschillen aangetoond werden, en ook dat deze verschillen in de loop der jaren steeds groter werden (minder verwijzingen naar speciaal onderwijs, minder zittenblijven). Pogingen om de variatie in de afzonderlijke projectuitkomsten te verklaren, liepen op niets uit: "All of these programs were effective, and none were more or less effective than others as measured by our school outcome indicators" (p. 462). Derhalve luidde de conclusie dat vroege interventies, ongeacht de vormgeving van het interventieprogramma, een gunstig effect hebben op het IQ (alleen een onmiddellijk effect) en op schoolloopbaanvariabelen (verwijzingen, zittenblijven). In deze conclusie komt onvoldoende tot uiting dat de gegevens en de analyses een tamelijk gefragmenteerde indruk maken. De terugloop in proefpersonen is aanzienlijk, de meeste vergelijkingen zijn gebaseerd op slechts een deel van de studies, de studies verschillen onderling aanzienlijk in de aard van hun onafhankelijke variabele en in evaluatie-opzet. Wij hebben deze review daarom op het vierde criterium (conclusies) als onvoldoende beoordeeld.

Het hoofdstuk van *Clarke-Stewart en Fein* (1983) in het 'Handbook of Child Psychology' biedt een breed perspectief op voorschoolse stimuleringsprogramma's. De review van onderzoek naar de effecten van zulke programma's bestrijkt slechts een gedeelte van het hoofdstuk. In dit gedeelte worden de effecten onderverdeeld naar vier domeinen. Wij bespreken hier alleen het effectonderzoek binnen het sociale en het intellectuele domein. Aan dit deel ligt een grote hoeveelheid onderzoek ten grondslag. Binnen het sociale domein worden circa 45 verschillende literatuurverwijzingen gegeven, binnen het intellectuele domein gaat het om ongeveer 35 verwijzingen. Het aantal onderliggende studies is niet precies te bepalen, omdat het onduidelijk is of elke verwijzing correspondeert met een afzonderlijke effectstudie. De auteurs verstrekken geen toelichting op de criteria die aan de keuze van de gereviewde studies ten grondslag hebben gelegen. De review is kwalitatief van aard. Dat wil zeggen, bij het coderen van de afzonderlijke studies en de integratie zijn geen objectieve of kwantitatieve procedures gevolgd. Slechts incidenteel wordt melding gemaakt van de significantie of van de grootte van de gevonden effecten. Voor het sociale domein luidt de hoofdconclusie dat kinderen na afloop van een programma sociaal competent of rijper zijn. Onduidelijk is of deze conclusie een referentiegroep veronderstelt van kinderen die géén programma gevolgd hebben. Voor het intellectuele domein wordt geconcludeerd dat de programma's een consistent positief effect laten zien dat meestal kleiner is dan één standaarddeviatie. Doordat de auteurs zo weinig inzicht geven in de aard van het onderliggende onderzoek en in de door hen gevolgde integratieprocedure, valt de validiteit van deze conclusies niet na te gaan. Wij hebben deze review op alle vier de criteria als onvoldoende gekwalificeerd.

McKey et al. (1985) is een review van al het tussen 1967 en 1982 verrichte onderzoek – gepubliceerd en ongepubliceerd – naar effecten van Head Start programma's. De aandacht ging uit naar effecten op kinderen (wat betreft hun cognitieve en sociaal-emotionele ontwikkeling en hun gezondheid), op gezinnen en op de lokale gemeenschappen waarbinnen de programma's geïmplementeerd werden. Wij be-

spreken hier alleen het deel dat betrekking heeft op de effecten op de kinderen zelf. De speurtocht naar relevant onderzoek is uitputtend geweest. De gehanteerde selectie-criteria en de coderingsprocedures worden uitgebreid verantwoord en maken een degelijke indruk. Van de 210 opgespoorde studies bleken er slechts 76 voldoende gegevens te bevatten voor een meta-analyse. Ter verklaring van de variatie in de gerapporteerde effectgroottes werd een uitgebreide set predictoren in stelling gebracht, waaronder het type onderzoeksoptzet ('pre-post' versus 'control-treatment'), kenmerken van de programma's en van de deelnemende kinderen. De conclusie is dat kinderen door Head Start een onmiddellijk voordeel ondervinden op alle drie de typen effectmaten (cognitie, sociaal-emotionele ontwikkeling, gezondheid). De gemiddelde effectgrootte binnen het cognitieve domein bedroeg .5 sigma, en voor het sociaal-emotionele domein .2 à .3 sigma. Evenwel, de effecten werden van jaar tot jaar kleiner. Twee tot drie jaar na afloop van het programma, bleek de winst voor Head Start kinderen geheel verdwenen te zijn. Voor deze laatste conclusie konden de reviewers overigens nog op slechts minder dan tien primaire studies terugvallen. Op alle vier kwaliteitscriteria hebben wij deze review een positieve score gegeven.

White (1985-1986) rapporteert over een database met meer dan 300 verschillende studies naar de effectiviteit van vroege-interventieprogramma's. Hiervan hebben er 99 betrekking op programma's voor achterstandskinderen, merendeels uitgevoerd in de jaren zestig en zeventig. Onze bespreking regardeert alleen dit deel van de review. Hoewel White volledigheid suggereert, heeft hij beduidend minder studies gereviewed dan McKey et al. (1985), die alleen onderzoek naar Head Start programma's (noodzakelijkerwijs na 1965) geanalyseerd hebben. Dit gegeven roept twijfel op aangaande de geclaimde volledigheid. Elke studie is op circa 90 kenmerken gecodeerd (onder meer wat betreft de proefpersonen, de interventie en de methodologische kwaliteit van het onderzoek). Desondanks is de review opvallend weinig specifiek aangaande de afhankelijke en onafhankelijke variabelen in de primaire studies. De auteur schat het onmiddellijke effect van de interventie op ongeveer .5

sigma, in de loop van drie tot vijf jaar na de interventie afnemend tot nul. In deze schatting zijn effecten op zeer verschillende afhankelijke variabelen samengevoegd. Ondanks het grote aantal gecodeerde variabelen per studie slaagt hij er niet in de variantie in de gerapporteerde effectgroottes te verklaren. Aangaande het soort programma dat het meest effectief is, levert de review feitelijk weinig op. Wij hebben deze review alleen op het vierde criterium, aangaande de correspondentie tussen de gepresenteerde feiten en de getrokken conclusie, positief gewaardeerd. Aangetekend zij hierbij, dat de negatieve beoordeling voornamelijk is ingegeven door de onvolledigheid van de rapportage wat betreft de gehanteerde zoek-, codeer- en integratieprocedures. Niet uit te sluiten is dat deze procedures op zichzelf beschouwd adequaat zijn uitgevoerd.

Stallings en Stipek (1986) Deze review bestrijkt het hele lager onderwijs. Wij bespreken hier alleen het deel dat betrekking heeft op voorschoolse stimulering. De auteurs bespreken alleen longitudinaal onderzoek. Methodologische overwegingen hebben een rol gespeeld bij de keuze van hun bronstudies (zie p. 728), maar omdat ze hun criteria niet prijs geven, is de kwaliteit van hun selectieprocedure niet te controleren. Aan één van de Consortium-studies wordt uitgebreide aandacht besteed, omdat "... the Perry Preschool Program stands out in having significant long-term findings ..." (p. 729). Door deze aanpak wordt bewust bijgedragen aan het verschijnsel 'publication bias' (een eenzijdige keuze van de te reviewen bronnen). De review is kwalitatief van aard, waarbij de resultaten op een niet-statistische wijze worden geïntegreerd. De reviewers laten zich niet uit over de grootte van de gerapporteerde effecten. In het slotwoord worden allereerst de conclusies van het 'Consortium for Longitudinal Studies' (1983) met instemming aangehaald (p. 740). Er wordt aan toegevoegd dat activiteiten om het gezin bij school te betrekken een gunstig effect hebben op de slaagkansen van kinderen. Deze laatste conclusie berust slechts op twee studies die een selecte keuze vormen uit een veel grotere groep studies naar het effect van ouderparticipatie. Een andere conclusie over het gunstige effect van Montessori-programma's op de cognitieve ontwikkeling berust op slechts één studie,

waarin niet duidelijk is geworden waarom dit effect wel voor vierjarige jongens maar niet voor vierjarige meisjes is gevonden. Deze review vertoont op alle vier criteria serieuze tekortkomingen.

Goldring en Presbey (1986). Centrale vraag van deze review is hoe sterk het effect van preschool programma's is op vier met name genoemde afhankelijke variabelen: IQ, lees- en rekenprestatie ('in grade 4') en het percentage leerlingen met een vertraagde schoolloopbaan. Als bronstudies heeft men de acht Consortium studies gekozen die relevante data bevatten. De effectgroottes van deze studies worden afzonderlijk gerapporteerd. Voor bijvoorbeeld IQ variëren ze van -.19 tot en met .86. Deze variatie is aanzienlijk, maar de auteurs concluderen desalniettemin dat de studies een homogeen positief effect laten zien (algemene effectgrootte voor IQ, vier tot zes jaar na de interventie: .42). Zij gaan voorbij aan grote verschillen in de aard van de voorschoolse stimuleringsprogramma's en ze doen geen pogingen de variatie in effectgrootte te verklaren. Het weglekken van de effecten onder invloed van de tijd wordt door de auteurs met statistische argumenten ontkend, maar hun figuren laten tegelijkertijd zien dat de 95-procents betrouwbaarheidsintervallen de nul insluiten. Hun conclusies stemmen niet overeen met de gepresenteerde feiten. Deze review schiet tekort wat betreft drie van de vier criteria.

Karweit (1989) bespreekt de korte- en lange-termijneffecten van voorschoolse stimuleringsprogramma's. Wat betreft volledigheid legt de auteur een sterke claim: "This chapter reviews the major literature on the effects of preschool." (p. 76). Maar in feite bespreekt ze slechts vijf primaire bronnen (waarvan er vier ook door het 'Consortium for Longitudinal Studies' besproken zijn, namelijk Gray's 'Early training project', het 'Perry Preschool Project' van Schweinhart en Weikart, het Louisville experiment van Miller en Bizzel en het onderzoek van Karnes et al.) en daarnaast de meta-analyse van McKey et al. (1985). In de review staan twee soorten vergelijkingen centraal. Eerst wordt - op basis van drie studies - het effect van wel of geen 'preschool' programma besproken. Van twee van de drie studies worden effectgroottes vermeld, van de derde alleen de richting van de verschillen.

Karweits conclusie is dat voorschoolse stimulering tot een onmiddellijk positief effect leidt ter grootte van ongeveer 1.00 sigma, afnemend tot ongeveer .20 sigma aan het eind van 'grade 3'. Deze conclusie is niet in overeenstemming met de vermelde gegevens. De genoemde effectgrootte van 1.00 is in feite een weergave van het grootste gerapporteerde effect, en derhalve niet representatief voor de door Karweit aangedragen effectgroottes. In het tweede deel van de review worden de effecten van programma's onderling vergeleken. Daarbij komen geen consistente verschillen in effectiviteit aan het licht. De analyse van de verschillen tussen de programma's is erg oppervlakkig. Er is geen poging gedaan de verschillen op vaste of systematische punten te beschrijven. Karweits conclusie luidt dat de programma's elkaar in effectiviteit weinig ontkomen en dat ze allemaal de moeite waard lijken. Door onduidelijkheden in de vermelde gegevens is de validiteit van deze conclusie niet na te gaan. Wij zijn van oordeel dat haar review op alle vier criteria tekort schiet.

De review van Farran (1990) bestaat uit twee delen, een voor 'disadvantaged' en een voor 'handicapped children'. Wij bespreken alleen het eerste deel, waarin 24 bronstudies worden samengevat. Omdat bij sommige studies meer interventieprogramma's betrokken waren, is het aantal gereviewde programma's iets groter, namelijk 32. De auteur laat zich niet uit over de totstandkoming van zijn selectie. Daarom is niet na te gaan waarom er bijvoorbeeld van de elf oorspronkelijke Consortiumstudies door Farran slechts zeven besproken worden. De integratie van de uitkomsten is kwalitatief van aard, waardoor geen inzicht gegeven wordt in de grootte van de gerealiseerde effecten, noch in de kenmerken van de interventies of van de studies die de effectgroottes zouden kunnen beïnvloeden. De auteur concludeert dat zorgvuldig geplande en uitgevoerde vroege-interventieprogramma's een reeks gunstige effecten kunnen hebben ('a range of potential benefits'), maar dat interventies gecontinueerd moeten worden om uitdoving te voorkomen. In feite, zo vervolgt Farran, hebben achterstandsléerlingen meer aan een verbetering van het onderwijs in de eerste vier jaren van de basisschool dan aan vroege-interventieprogramma's. De door hem bespro-

ken studies leveren geen bevindingen op die voor deze slotconclusie relevant zijn. Naar ons oordeel voldoet Farrans review aan geen van de vier criteria.

White, Taylor en Moss (1992) reviewen een veel breder probleemgebied dan wij hier behandelen, namelijk de vraag of ouderparticipatie het effect van vroege-interventieprogramma's gunstig beïnvloedt. In een apart onderdeel, het enige deel dat wij hier aan de orde stellen, beantwoorden ze de vraag of interventieprogramma's mét een ouder-component sterkere effecten teweegbrengen dan programma's zonder zo'n component. Uit een in het begin van de jaren tachtig aangelegd bestand van onderzoeken (met ruim 350 studies) werden alle studies geselecteerd die voor deze vraag relevant materiaal bevatten. Vermoedelijk betreft het hier dezelfde studies die ook door de eerder besproken White (1985-1986) gereviewed zijn. Elke studie is op vele tientallen aspecten gecodeerd, onder meer wat betreft de afhankelijke en onafhankelijke variabelen en de kwaliteit van de onderzoeksopzet. Als uitkomsten van de studies werden effectgroottes bekeken. Voor details over de gehanteerde procedures wordt naar tussentijdse rapporten verwezen. Wij zijn er niet in geslaagd deze rapporten te bemachtigen, zodat we slechts kunnen constateren dat de beschikbare rapportage onvoldoende gegevens bevat om de kwaliteit van de coderingen en van de gegevensintegratie te beoordelen. Beperken we ons tot de effectgroottes voor studies met achterstandskinderen waarbij een 'center-based' aanpak werd gehanteerd, dan bedroegen deze gemiddeld .52, respectievelijk .53 voor de studies met, respectievelijk zonder oudercomponent. Deze cijfers hebben alleen betrekking op de 43 studies met een hoge interne validiteit. De 54 studies met een lage interne validiteit lieten wel een verschil zien. De conclusie van de auteurs luidt dat programma's mét een oudercomponent géén sterkere effecten laten zien dan de programma's die zo'n component missen. Deze conclusie blijft dicht bij de feiten zoals in het artikel weergegeven. Gezien de overeenkomsten met de review van White (1985-1986) hebben wij dezelfde oordelen gegeven.

Sylva en Wiltshire (1993). Centrale vraag in deze review is of voorschoolse stimulerings-

programma's een blijvend effect hebben op de latere levensloop van kinderen ('a lasting effect on the course of children's later lives'). De auteurs geven aan alleen de 'key studies' te bespreken, maar weiden over hun selectie verder niet uit. Geconstateerd moet worden dat slechts één van de besproken studies bij hun vraagstelling past (namelijk het Perry Preschool Project). Het merendeel van de studies evalueert alleen effecten op korte termijn en binnen het schoolse domein. De bespreking van de afzonderlijke studies is onsystematisch en niet kwantitatief. In sommige gevallen worden effectgroottes vermeld, in andere gevallen niet. De auteurs concluderen dat voorschoolse interventie op de korte termijn tot meetbare effecten leidt binnen het schoolse en sociale domein. Voorts, als de interventie van hoge kwaliteit is, zijn de effecten op het cognitieve en sociale vlak van blijvende aard. De effecten gelden voor kinderen uit alle sociale groepen, maar zijn het sterkst voor achterstandsleerlingen. Geen van deze conclusies wordt gestaafd door het besproken onderzoek. Hetzelfde geldt voor de andere door Sylva en Wiltshire getrokken conclusies betreffende onder meer het economische voordeel van 'high quality early education'. Op alle vier criteria waarden wij deze review negatief.

4 Conclusies en discussie

Wij beginnen de afsluitende paragraaf met de conclusie ten aanzien van de secundaire vraagstelling, de vraag naar de kwaliteit van de besproken reviews. Onze conclusie is dat de kwaliteit van de reviews over het geheel genomen teleurstellend laag is. Slechts twee van de tien reviews, die van het 'Consortium for Longitudinal Studies' (1983) en van McKey et al. (1985), hebben wij op drie of vier van onze criteria van een positief oordeel kunnen voorzien. De resterende acht reviews zijn op nagenoeg alle criteria als onvoldoende gekwalificeerd. Wij vinden dit een onthutsende conclusie, mede omdat de bekritiseerde reviews in gerenommeerde tijdschriften of handboeken verschenen zijn. In het discussiedeel van deze paragraaf komen we op deze kwestie nog terug.

Bij onze conclusie ten aanzien van de primaire vraagstelling betrekken we slechts de re-

views die de toets van onze kritiek kunnen doorstaan. Dat zijn er, zoals gezegd, maar twee. Wat valt er ten aanzien van de effectiviteit van voorschoolse stimuleringsprogramma's voor lering te trekken? We onderscheiden vijf deelconclusies.

Ten eerste, direct na afloop van de interventie laten IQ-testen een gemiddelde winst van ongeveer .5 sigma zien, zijnde ongeveer 7 IQ-punten. Deze winst verdwijnt betrekkelijk snel ('Consortium for Longitudinal Studies', 1983: na het eerste jaar is de winst ongeveer gehalveerd; McKey et al., 1985: na één jaar is de winst nagenoeg helemaal verdwenen).

Ten tweede, wat betreft de vorderingen binnen het schoolse domein (schoolrijpheid, lezen, rekenen) blijkt alleen uit de review van McKey et al. (1985) een onmiddellijke winst van ongeveer .5 sigma. Deze winst ebt in ongeveer twee jaar weg. De review van het 'Consortium for Longitudinal Studies' (1983) laat zien dat een onmiddellijke winst op dit vlak soms wel, maar soms ook niet aantoonbaar is. De overheersende tendens is dat een eventueel effect op dit domein zich hoogstens tot en met 'grade 1' (groep 3 in ons basisonderwijs) uitstrekt.

Ten derde, gunstige effecten op de schoolloopbaan van de leerlingen (wat betreft verwijzingen naar speciaal onderwijs en zittenblijven) worden in slechts enkele studies aangetoond. Dergelijke effecten zijn derhalve omstreden, mede omdat het niet plausibel is dat leerlingen een gunstiger schoolloopbaan zouden weten te realiseren, zonder de daarbij behorende winst op het vlak van de schoolvorderingen.

Ten vierde, het onmiddellijke effect binnen het sociaal-emotionele domein varieert volgens McKey et al. (1985) van .4 sigma (namelijk voor sociaal gedrag) tot .2 sigma (namelijk voor zelfvertrouwen en prestatie-motivatie). De winst verdwijnt na het tweede jaar. Het 'Consortium for Longitudinal Studies' (1983) heeft binnen dit domein geen metingen gerapporteerd.

Ten vijfde, het is duidelijk dat de stimuleringsprogramma's ten aanzien van hun doelstellingen en werkwijze enorm van elkaar verschillen. Maar de consequenties van deze verschillen voor de effectiviteit van de programma's zijn nagenoeg onbekend. Het 'Con-

sortium for Longitudinal Studies' (1983) suggereert dat alle programma's even effectief zijn, maar deze conclusie wordt niet door de aangevoerde feiten gestaafd. De review van McKey et al. (1985) levert als enige houvast dat programma's met een sterk gestructureerde, schoolse gerichtheid effectiever zijn dan programma's die een dergelijke gerichtheid missen.

In het discussiegedeelte van deze paragraaf stellen we drie onderwerpen aan de orde.

Het eerste onderwerp is het belang van de aangetoonde effecten. Op zichzelf lijken effecten ter grootte van een halve sigma natuurlijk zeer de moeite waard. Maar de wetenschap dat de effecten in betrekkelijk korte tijd weglekken, zou aanleiding kunnen geven tot een geringschattende houding. Wij menen dat zo'n houding niet terecht is, zolang niet duidelijk is wat de achtergrond van dit weglekken is. Het is opvallend dat de besproken reviews hierover nauwelijks theoretiseren. Zigler, Styfco & Gilman (1993, p. 21) brengen naar voren dat de ontwikkeling van kinderen - ná de voorschoolse periode - vanzelfsprekend van veel meer grootheden afhankelijk is dan het éénjarige voorschoolse programma wat ze al dan niet gevolgd hebben. Invloeden van school en huis overschaduwden al snel de winst die kinderen als vierjarige ten gevolge van een voorschoolse programma geboekt hebben. Steeds meer wint het inzicht veld dat vroege interventie alleen niet genoeg is om tekorten in de thuisomgeving te verhelpen. Zigler (1987) verwoordt dit als volgt: "We simply cannot inoculate children in one year against the ravages of a life of deprivation" (p. 258). Nodig zijn vervolprogramma's die een ingebouwde compensatie bieden voor kinderen die thuis niet optimaal in hun ontwikkeling ondersteund worden. Longitudinaal onderzoek naar de effecten van voorschoolse stimuleringsprogramma's - zonder dat aan kinderen zulke vervolprogramma's worden aangeboden - vormt een miskenning van ontwikkelingspsychologische inzichten. Het weglekken van de winst pleit in feite meer voor het ontwerpen en evalueren van vervolprogramma's dan tot een geringschatting van het belang van voorschoolse stimuleringsprogramma's.

De relevantie van deze vijf bevindingen

voor het Nederlandse onderwijsbeleid - het tweede discussieonderwerp - vormen feitelijk niet het onderwerp van dit artikel. Toch achten wij het raadzaam in dit verband een waarschuwend woord te laten horen. De hier beschreven bevindingen hebben betrekking op interventies, exclusief gericht op achterstandskinderen in de Verenigde Staten en voor het merendeel uitgevoerd tussen 1960 en 1980. Tussen 'toen' in de Verenigde Staten en 'nu' in Nederland zijn belangrijke verschillen die de generaliseerbaarheid ernstig kunnen beperken. We wijzen erop dat Nederlandse kleuters twee tot drie jaar onderwijs volgen, voordat ze in groep 3 met het meer formele taal- en rekenonderwijs starten. In de Verenigde Staten was - en is - de periode waarin kinderen geleidelijk in de eisen van het schoolsysteem kunnen ingroeien, veelal slechts één schooljaar, waarin bovendien veel kinderen slechts een halve-dagsprogramma volgen. Ook in andere opzichten zijn er verschillen. Het is onze indruk dat de kwaliteit van de thuisopvoeding van Nederlandse achterstandsleerlingen veel minder achterblijft ten opzichte van de vorming en educatie die leerlingen op school krijgen, dan van Amerikaanse achterstandskinderen. Als deze indruk juist zou zijn, zouden Nederlandse voorschoolse interventieprogramma's - zelfs als deze op de zwaarste achterstandscategorieën gericht zouden zijn - tot minder grote effecten leiden dan in de hier gereviewde literatuur gerapporteerd worden.

Rest ons de discussie van het derde onderwerp: hoe komt het dat hier de conclusie getrokken wordt dat zó veel van de besproken reviews in zo ernstige mate tekort schieten? Het antwoord zou kunnen zijn, dat wij 'unfaire' criteria hebben aangelegd. Zo zou men tegen kunnen werpen dat onze eisen te veel gemodelleerd zijn op een kwantitatieve reviewprocedure, ook wel aangeduid als meta-analyse ter onderscheiding van een traditionele review waarin de integratie voornamelijk kwalitatief van aard is. Het is inderdaad waar dat aan de hier gehanteerde criteria in feite alleen door meta-analyses voldaan kan worden. Maar bij deze tegenwerping passen twee argumenten. In de eerste plaats is de hoeveelheid te reviewen onderzoek zo groot, dat een voornamelijk kwalitatieve werkwijze omslachtig is en niet meer kan voldoen aan de reproduceerbaarheidseis.

In de tweede plaats is het gezien de publikatie-data van de reviews ook werkelijk 'fair' de meta-analytische criteria hier toe te passen. Wij wijzen erop dat alle reviews gepubliceerd zijn, nadat de meta-analytische werkwijze ontwikkeld is. De fundamenteen voor meta-analyse zijn immers al in de jaren zeventig gelegd door onder andere Glass (1976) en Rosenthal (1978). Wij houden derhalve staande dat onze criteria afgeleid zijn uit zeer algemene wetenschappelijke eisen aangaande objectiviteit en reproduceerbaarheid. Wij eisen, met andere woorden, dat denkbeeldige andere reviewers - gegeven eenzelfde vraagstelling en gegeven eenzelfde 'body of research' - tot dezelfde conclusies zouden moeten komen. Als aan deze eis niet tegemoet gekomen wordt, ontvalt de eigenlijke zin aan het wetenschapsbedrijf.

Kan het zijn dat reviewers op zich hun werk goed doen, maar dat redacties van tijdschriften en handboeken het hun onmogelijk maken om verantwoord verslag te doen? Men kan opmerken dat de beide 'goedgekeurde' reviews over veel rapportageruimte beschikten, terwijl de andere reviews zich moesten voegen binnen de beperkingen van een artikel in een tijdschrift of van een hoofdstuk in een handboek. Toch verworpen wij voor de meeste van de afgekeurde reviews ook deze mogelijkheid. Slechts bij de reviews van White (1985-1986) en van White, Taylor en Moss (1992) hebben wij de mogelijkheid opengelaten, dat bestudering van onderliggende rapporten - waarover we in die gevallen niet konden beschikken - alsnog tot een positieve waardering zou kunnen leiden.

Aan de conclusie dat het merendeel van de reviews onbetrouwbaar is, valt ons inziens niet te ontkomen. De punten waarop de reviews tekort schieten, betreffen alle vier hier onderscheiden criteria.

Aangaande het criterium *selectie* treden twee problemen aan het licht. Opvallend is dat reviewers weinig aandacht besteden aan een zorgvuldige domeinafgrenzing vooraf, vóórdat het zoekproces wordt gestart. Die afgrenzing zou voort moeten komen uit een explicitering van de onderzoeksvraag. Diverse reviewers komen wat betreft hun vraagstelling niet verder dan een terloopse opmerking dat zij interventieprojecten reviewen. Specificaties die achterwege blijven, betreffen onder meer een type-ring van de aard van de afhankelijke

variabelen, de aard en de leeftijd van de leerlingpopulaties waarop de interventies zich richten en het jaar waarin de interventies geëvalueerd zijn. Weliswaar bevatten de meeste reviews op deze punten een specificatie achteraf. Maar zo'n specificatie achteraf kan een specificatie vooraf niet vervangen, zo min als een beschrijving van het resultaat van een steekproeftrekking in de plaats kan treden van een specificatie van de populatie. Ook de gevolgde zoekprocedures - dat is het tweede probleem - blijven onderbelicht. Gevolg is dat de meeste reviewers geen garantie kunnen geven omtrent de volledigheid of representativiteit van het onderzoek waarop zij zich baseren.

Ten aanzien van het criterium *coding* is het opvallend dat in de meeste reviews geen kenmerken gecodeerd zijn, noch kenmerken betreffende de afhankelijke en onafhankelijke variabelen of de aard van de populatie, noch kenmerken die de publikatiestatus betreffen. Zo zijn de reviews in feite heel weinig informatief over de aard van de interventies waarop de onderzoeken betrekking hebben. Effecten van waarschijnlijk zeer uiteenlopende programma's worden zonder meer bij elkaar gevoegd. Toch is aannemelijk dat de vormgeving van een programma, met aspecten als de doelstellingen, de leertheoretische principes, de intensiteit en de duur, in sterke mate de effectiviteit bepalen. Wat we in feite constateren, is dat de reviews arm aan theorie zijn. Immers, theorie-gestuurde uitspraken over hoe jonge kinderen zich ontwikkelen dicteren idealiter welke variabelen er ten minste gecodeerd dienen te worden.

Wat betreft het criterium *integratie* is de grootste tekortkoming dat de uitkomsten van de studies hoofdzakelijk in kwalitatieve zin worden besproken. Effectgroottes en significantieniveaus spelen een geringe rol. Daardoor worden variaties in de uitkomsten van de afzonderlijke studies niet zichtbaar en vervalt de noodzaak voor deze variatie een adequate verklaring te postuleren.

Ook bij het laatste criterium, *conclusies*, zijn wij tot negatieve evaluaties gekomen. In een aantal reviews worden conclusies niet expliciet getrokken, maar alleen impliciet verwoord met weglating van relevante condities. Soms ook worden conclusies overgeslagen en stappen reviewers van de resultaten direct over

naar de implicaties voor de praktijk van de interventies. In meer dan één geval hebben wij geconstateerd dat conclusies niet in overeenstemming zijn met de gepresenteerde data. Het lijkt, kortom, of sommige reviewers subjectieve opvattingen als uitkomsten van een wetenschappelijk verantwoorde review presenteren.

Noot

1. De opdracht is verstrekt door het Instituut voor Onderzoek van het Onderwijs (SVO), project 94807.

Literatuur

(De besproken reviews zijn met een asterisk aangeduid.)

- Clark, M. M. (1988). *Children under five: educational research and evidence*. New York: Gordon and Breach Science Publishers.
- *Clarke-Stewart, K. A., & Fein, G. G. (1983). Early childhood programs. In P. H. Mussen (Ed.), *Handbook of Child Psychology, formerly Carmichael's Manual of child psychology; Volume II: Infancy and developmental psychobiology* (pp. 917-1000). New York: Wiley.
- *Consortium for Longitudinal Studies, The (1983). *As the twig is bent ...; Lasting effects of preschool programs*. Hillsdale, NJ: Lawrence Erlbaum.
- *Farran, D. C. (1990). Effects of intervention with disadvantaged and disabled children: A decade review. In S. J. Meissels & J. P. Shonkoff, *Handbook of early childhood intervention* (pp. 501-539). Cambridge: Cambridge University Press.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8.
- *Goldring, E. B., & Presbey, L. S. (1986). Evaluating preschool programs: a meta-analytic approach. *Educational Evaluation and Policy Analysis*, 8, 179-188.
- Harrell, A. (1983). *The effects of Head Start on children's cognitive development: preliminary report* (Publication No. (OHDS) 83-31186). Washington, DC: DHHS Publications.
- *Karweit, N. (1989). Effective preschool programs for students at risk. In R. E. Slavin, N. L. Karweit & N. A. Madden, *Effective programs for students at risk* (pp. 75-102). Needham Heights: Allyn & Bacon.
- Kulik, J. A., & Kulik, C.-L. C. (1989). Meta-analysis in education [special issue]. *International Journal of Educational Research*, 13(3).
- Lazar, I., & Darlington, R. (1982). The lasting effects of early education: a report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development*, 47 (nos. 2-3, serial no. 195).
- *McKey, H. R., Condelli, L., Ganson, H., Barrett, B., McConkey, C., & Plantz, M. (1985). *The impact of Head Start on children, families and communities*. Washington, DC: DHHS publication.
- Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin*, 85, 185-193.
- Seifert, K. L., & Hoffnang, R. J. (1994). *Child and adolescent development* (third edition). Boston: Houghton Mifflin Company.
- Slavin, R. E. (1984, october). Meta-analysis in education: How has it been used? *Educational Researcher*, 13, 6-15.
- *Stallings, J. A., & Stipek, D. (1986). Research on early childhood and elementary school teaching programs. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd edition) (pp. 727-753). New York: MacMillan.
- *Sylva, K., & Wiltshire, J. (1993). The impact of early learning on children's later development: A review prepared for the RSA inquiry 'Start Right'. *European Early Childhood Education Research Journal*, 1, 17-40.
- Westinghouse Learning Corporation (1969). *The impact of Head Start: an evaluation of the effects of Head Start on children's cognitive and affective development*. ERIC (ED 036321).
- *White, K. R. (1985-1986). Efficacy of early intervention. *The Journal of Special Education*, 19, 401-416.
- *White, K. R., Taylor, M. J., & Moss, V. D. (1992). Does research support claims about the benefits of involving parents in early intervention programs? *Review of Educational Research*, 62, 91-125.
- Zigler, E. F. (1987). Formal schooling for four-year-olds? No. *American Psychologist*, 33, 254-260.
- Zigler, E., Styfco, S. J., & Gilman, E. (1993). The national Head Start program for disadvantaged preschoolers. In E. Zigler & S. J. Styfco (Eds.), *Head Start and beyond; a national plan for extended childhood intervention* (pp. 1-42). New Haven: Yale University Press.

Auteurs

H. Blok en P.P.M. Leseman zijn verbonden aan het SCO-Kohnstamm Instituut van de Faculteit Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam

Adres: Wibautstraat 4, 1091 GM Amsterdam

Abstract

Effects of preschool programs: a review of reviews

H. Blok, P.P.M. Leseman. *Pedagogische Studiën*, 1996, 73, 184-197.

This article deals with the question how preschool programs affect children's development. Ten reviews, published between 1983 and 1993, have been evaluated on four criteria: 1. the selection of the source studies; 2. the coding of the studies' characteristics; 3. the way the results were integrated and 4. the correspondence between the results and the conclusions. Eight out of these ten studies proved to be seriously defective with regard to two or more of the criteria. The two remaining reviews permit the following conclusions: a. IQ tests and achievement tests show an immediate gain of about half a sigma; b. the benefit on the social domain is about a quarter of a sigma; c. in most cases the benefit is no longer demonstrable after one to two years; d. only a few studies show an effect on the school career; e. little is known about the characteristics of the programs which affect the program's efficiency. Since most studies were carried out among American disadvantaged children between 1960 and 1980, it is difficult to generalize these conclusions to the present Dutch educational system.

Coördinatie van het vakonderwijs binnen scholen voor voortgezet onderwijs: vaksecties en schoolleiding

B. Witziers

Samenvatting

In de schoolorganisatieliteratuur worden scholen veelal voorgesteld als nauwelijks geïntegreerde organisaties. Er wordt gewezen op het ontbreken van elementen als onderwijskundig leiderschap, overeenstemming over onderwijsdoelstellingen, onderwijsbeleid en samenwerking tussen docenten.

Voor onderzoekers houdt dit beeld de opgave in om de juistheid ervan empirisch te toetsen. In onderhavig onderzoek is dit gebeurd door de coördinerende rol van vaksecties en het (midden)management ten aanzien van het vakonderwijs in scholen voor algemeen voortgezet onderwijs nader te onderzoeken. Uit de resultaten van dit onderzoek blijkt dat secties hun vakonderwijs sterker coördineren dan op basis van de literatuur verwacht mag worden, al zijn er tussen secties op verschillende vakgebieden wel verschillen waar te nemen. Meer passend met de literatuur is het feit dat de schoolleiding en vaksectieleiders wat betreft coördinatie van het vakonderwijs een geringe rol spelen. Verder laten de resultaten zien dat er naar coördinerend vermogen duidelijk te onderscheiden schooltypen zijn. Aan de ene kant van het spectrum zijn er scholen waar de coördinerende rol van vaksecties en (midden)management zwak ontwikkeld is, aan de andere kant van het spectrum scholen met tegengestelde kenmerken.

1 Inleiding

Coördinatie binnen scholen voor algemeen voortgezet onderwijs is het thema van het hier te presenteren onderzoek. Coördinatie is een organisatiekundig begrip en heeft betrekking op de afstemming en ordening van de taken van de organisatieleden. Integratie is daarbij het doel, coördinatie het middel om dat te verwezenlijken (Mintzberg, 1979; Pelkmans &

Vrieze, 1987).

Over coördinatie binnen scholen is in de schoolorganisatieliteratuur veel geschreven. Grofweg is daarbij een tweetal benaderingen te onderscheiden. De eerste benadering benadrukt dat scholen nauwelijks geïntegreerde organisaties zijn. Het meest duidelijk komt dit naar voren in begrippen als de school als losjes gekoppeld systeem (Meyer & Rowan, 1983; Weick, 1976) en de school als vuilnisvatorganisatie (Cohen, March & Olsen, 1972). Deze begrippen verwijzen naar een situatie binnen scholen, waarbij individuen weliswaar van dezelfde bronnen gebruik maken, maar verder weinig met elkaar te maken hebben. Het ontbreken van integrerende elementen als onderwijskundig leiderschap, overleg en samenwerking tussen docenten, onderwijsbeleidsvorming en consensus ten aanzien van doelen en inhouden van het onderwijs kunnen daarbij voor deze situatie verantwoordelijk gesteld worden. Een verwant beeld komt naar voren in die modellen waarin het professioneel-bureaucratische karakter van schoolorganisaties benadrukt wordt. Van belang zijn hier de school als professionele bureaucratie (Mintzberg, 1979) en het 'Interacting Spheres Model' (Hanson, 1979). Ook in deze modellen wordt de structurele 'losheid' van de schoolorganisatie benadrukt, met name vanwege het accent dat gelegd wordt op de (semi-)professionele autonomie van de docent. Dit kenmerk zorgt ervoor dat coördinatie problematisch is. Ten eerste vanwege het feit dat docenten ieder in hoge mate hun eigen werk beheren en controleren, ten tweede omdat de rol van de schoolleiding (grotendeels) beperkt blijft tot het beheersmatig terrein. Hoezeer deze benadering het denken over schoolorganisaties domineert kan opgemaakt worden uit tal van artikelen. Zo wijzen Boyd en Crowson (1981), Willower (1979) en Ingersoll (1994) op de overheersende plaats van het model van de school als losjes

gekoppeld systeem in het denken over schoolorganisaties in de Verenigde Staten, terwijl in Nederland deze rol, getuige leerboeken, met name weggelegd lijkt voor het model van de professionele bureaucratie (Meijnen, 1987; Scheerens & Van Vilsteren, 1988; Smets, 1985a; Van Vilsteren, 1984; Van Vilsteren & Visscher, 1987).

Verschillende auteurs hebben er op gewezen dat de bovenstaande benadering een statisch beeld van scholen inhoudt (Schuit & Slegers, 1990; Witziers, 1992). Met andere woorden, alle scholen worden geacht weinig geïntegreerde organisaties te zijn. De tweede benadering nu breekt met deze veronderstelling en wijst erop dat scholen naar coördinatie zullen verschillen. Als zodanig komt een dergelijk standpunt bijvoorbeeld naar voren in beschouwingen over effectieve scholen, waarin met een zekere regelmaat een onderscheid wordt gemaakt tussen losjes gekoppelde/niet-effectieve scholen enerzijds en stevig-gekoppeld/wel-effectieve scholen anderzijds (Edmonds, 1979; Rosenholtz, 1985). Meer expliciet komen deze verschillen tussen scholen naar voren in de schoolorganisatiemodellen van Marx (1975). Zo stelt Van der Krogt (1983) dat de markantste verschillen tussen de door Marx onderscheiden modellen bestaan in de wijzen waarop coördinatie gestalte krijgt. Zo is van coördinatie nauwelijks sprake in het segmentale model, terwijl het tegenovergestelde het geval is in het coöperatief-gereguleerde model. Belangrijkste verklarende factoren voor deze verschillen zijn, althans oorspronkelijk, veranderingen in de doelstellingen en vormgeving van het primaire proces, terwijl in latere publikaties daarnaast de nadruk komt te liggen op factoren als de omvang van de school, de cultuur binnen de school en de rol van de schoolleiding (Marx, 1984, 1987).

Ten aanzien van beide benaderingen kan de vraag gesteld worden of zij een accuraat beeld geven van de wijze waarop schoolorganisaties functioneren. Zo kan wat betreft de eerste benadering de vraag gesteld worden of scholen in algemene zin inderdaad nog wel als weinig geïntegreerde organisaties getypeerd kunnen worden. Te wijzen valt in dit verband op tal van factoren binnen en rondom de school (fusies, schaalvergroting, toegenomen aandacht voor onderwijskundig leiderschap en beleidsvor-

ming, introductie basisvorming) die de wijze waarop scholen functioneren essentieel gewijzigd kunnen hebben. De tweede benadering roept de vraag op of de wijzen waarop verschillen tussen scholen geconceptualiseerd worden inderdaad een geschikt kader vormen om werkelijke verschillen in functioneren tussen scholen te vatten. In het bijzonder kan hier verwezen worden naar de discussie omtrent de empirische validiteit van de Marxse modellen (Van der Krogt, 1983; Scheerens, 1989).

2 Probleeminkadering

Coördinatie is binnen organisaties een veelomvattend verschijnsel. Dit geldt, getuige de literatuur, ook voor schoolorganisaties. Pelkmans en Vrieze (1987) stellen ten aanzien van schoolorganisaties dat coördinatie plaatsvindt op verschillende niveaus. Zij onderscheiden (a) externe coördinatie, bijvoorbeeld in de vorm van het vastleggen van de eindexameninhoud en -normen, (b) coördinatie op schoolniveau waar met name de schoolleiding een belangrijke rol speelt, en (c) coördinatie op het niveau van een organisatiedeel, bijvoorbeeld vaksecties. Daarnaast merken zij op dat coördinatie op verschillende domeinen plaats kan vinden. Zo kan coördinatie plaatsvinden op het beheersmatige terrein of op meer onderwijsinhoudelijke terreinen. Als zodanig vereist deze omvatendheid dat het onderzoeksterrein afgebakend wordt. In het hier te presenteren onderzoek gaat het primair om de wijze waarop vaksecties hun vakonderwijs coördineren. Daarnaast is in het onderzoek aandacht besteed aan de coördinerende rol van vaksectieleiders en de schoolleiding ten opzichte van het vakonderwijs.

Vanuit een perspectief waarin de school getypeerd wordt als losjes gekoppeld of als professioneel-bureaucratisch van aard mag verwacht worden dat vaksecties hun onderwijs niet of nauwelijks zullen coördineren. Vanuit deze modellen wordt immers de nadruk gelegd op verschijnselen als autonomie, gebrek aan samenwerking tussen docenten en conflicterende visies op het onderwijs. Alle aspecten die in hoge mate de ontwikkeling van vaksecties in de weg zullen staan. Een belangrijke vraag is echter of deze aanname ten opzichte van vaksecties terecht is. In de eerste plaats kan gewe-

zen worden op het feit dat vaksecties binnen scholen voor voortgezet onderwijs een 'normaal' verschijnsel zijn. Siskin (1991) stelt dat in de Verenigde Staten sinds de jaren dertig de opdeling van de schoolorganisatie in vaksecties de dominante wijze van organiseren is in scholen voor voortgezet onderwijs. Voor de Nederlandse situatie stellen Van Vilsteren en Witziers (1990a) dat sinds de jaren zestig, zowel gestimuleerd door de toegenomen schoolgrootte als door omvattende innovatie- en fusieprocessen, vaksecties een wezenlijk onderdeel uitmaken van scholen voor voortgezet onderwijs. In de tweede plaats kan het belang van vaksecties nader onderstreept worden door te wijzen op het feit dat het bij vaksecties gaat om een bundeling van specialisten die vakkennis en referentiekaders delen met alle consequenties van dien voor de interactiepatronen binnen schoolorganisaties. Deze beide aspecten geven aanleiding om te veronderstellen dat vaksecties hun onderwijs mogelijk-erwijs meer coördineren dan veelal wordt verondersteld. Duidelijk is dat een dergelijke bevinding het gangbare beeld van schoolorganisaties nuanceert. Eenzelfde redenering kan daarbij in principe gevolgd worden voor de andere onderzoeksobjecten, schoolleidingen en vaksectieleiders, zij het dat empirische resultaten, met name op het terrein van schoolleiderschap (Brandsma & Stoel, 1986; Van der Grift, 1987; Stoel, 1994), sterk in de richting wijzen van het feit dat zij inderdaad nauwelijks coördinerende taken ten opzichte van het onderwijs ter hand nemen.

Een tweede relevante vraag vloeit voort uit het feit dat in de literatuur verondersteld wordt dat scholen verschillen naar coördinerend vermogen. Het meest duidelijk komt dit naar voren in de door Marx ontwikkelde modellen. Eerder is op dit punt Van der Krogt (1983) aangehaald met de stelling dat de markantste verschillen tussen de door Marx onderscheiden modellen bestaan in de wijze waarop coördinatie gestalte krijgt. Meer concreet betekent dit dat er in toenemende mate verschillende elementen of onderdelen aanwezig zijn, waartussen bovendien meerdere samenhangen te constateren zijn. In de onderzoekspraktijk heeft deze stelling veelal geleid tot het afleiden van operationele kenmerken uit de Marxse modellen, waarna vervolgens getoetst is of de in de

modellen veronderstelde samenhang tussen kenmerken in de empirie terug te vinden is. Sommige onderzoekers (Van der Krogt, 1983; Mahieu, 1988) geven daarbij aan dat de modellen van Marx goed bruikbaar zijn op dit punt, terwijl de resultaten van ander onderzoek de empirische status van deze modellen niet duidelijk bevestigen (Van der Krogt & Oosting, 1990; Van Marwijk Kooy-von Baumhauer, 1984; Pelkmans & Vrieze, 1987; Schuit, 1993; Slegers, Bergen & Giesbers, 1992; Smets, 1985b).

Ondanks het feit dat in Nederland in de empirie blijkbaar moeilijk schooltypen te vinden zijn die in overeenstemming zijn met de theoretische modellen van Marx, blijft hier toch een belangrijke vraag liggen voor schoolorganisatieonderzoek. Voor dit onderzoek heeft deze constatering geleid tot de vraag of er naar de mate van coördinatie van het vakonderwijs schooltypen te vinden zijn die (in meer of mindere mate) overeenkomen met de Marxse modellen. Van belang is hier met name de specifieke rol die vaksecties en (midden)-management vervullen in de Marxse modellen met betrekking tot coördinatie van het vakonderwijs. Waar in het segmentale model het coördinerend vermogen van zowel vaksecties als (midden)management nauwelijks ontwikkeld is, is het tegenovergestelde het geval in het coöperatief gereguleerde model. Met andere woorden, de Marxse modellen wijzen er op dat scholen theoretisch te onderscheiden zijn naar de mate waarop vaksecties en midden(management) het onderwijs coördineren. Dit onderzoek probeert nu na te gaan of dergelijke theoretisch veronderstelde configuraties in de empirie terug te vinden zijn en zodoende een bijdrage te leveren aan de vraag of de Marxse modellen inderdaad een geschikt kader zijn om verschillen tussen scholen in kaart te brengen.

3 Het onderzoek

3.1 Vraagstellingen

Vanuit het bovenstaande kan afgeleid worden dat de kern van het onderzoek werd gevormd door een tweetal vragen. De eerste vraag heeft betrekking op de aard en mate van coördinatie van het vakonderwijs. Een antwoord op deze vraag wordt gezocht door in te gaan op de wijze

waarop vaksecties en het (midden)management binnen scholen voor algemeen voortgezet onderwijs het vakonderwijs coördineren.

De tweede vraag die centraal staat in dit onderzoek vloeit voort uit de door Marx veronderstelde verschillen tussen scholen wat betreft coördinerend vermogen. In het bovenstaande is gesteld dat op basis van de Marxse modellen scholen te typeren zijn op basis van de mate waarin coördinatie van het vakonderwijs door zowel vaksecties als (midden)management plaatsvindt. De specifieke onderzoeksvraag luidt nu of deze theoretisch veronderstelde typen inderdaad in de empirie terug te vinden zijn.

3.2 Onderzoeksofzet

Naar opzet kan het onderzoek getypeerd worden als een vergelijkend onderzoek. Een dergelijk onderzoek maakt in het algemeen gebruik van een momentopname binnen een groot aantal onderzoekseenheden en is daarmee geschikt voor het vaststellen en/of toetsen van verbanden en het verkrijgen van generaliseerbare onderzoeksgegevens (Van der Zwaan, 1984). Aangezien zowel het toetsen van verbanden als generalisatie tot de doelstellingen van dit onderzoek behoorden, lag de keuze voor een dergelijk onderzoekstype min of meer voor de hand. Een vergelijkend onderzoek is bovendien adequaat om op grond van empirische gegevens typologieën van scholen te construeren (Everitt, 1974).

3.3 De operationele steekproef

Bij een vergelijkend onderzoek wordt meestal gepoogd het aantal onderzoeksobjecten te maximaliseren. Duidelijk is echter dat gezien de heterogeniteit en omvang van de potentieel te onderzoeken populatie (alle scholen voor algemeen voortgezet onderwijs en alle secties daarbinnen) een nadere inperking noodzakelijk was. In concreto gaat het om de vraag welke vaksecties, docenten en schooltypen bij het onderzoek betrokken moesten worden.

Wat betreft de vaksecties is gekozen voor een drietal vakken, te weten wiskunde, Engels en geschiedenis. Verreweg de belangrijkste reden hiervoor is dat met de keuze voor deze vakken een drietal verschillende kennisdomeinen binnen de school bestreken werd. Uit de literatuur kan ontleend worden dat er tussen vaksec-

ties al naar gelang het kennisdomein waarop zij opereren de nodige verschillen in functioneren zijn te verwachten. (Ax, 1985; Hylkema, 1990; Murphy, 1988; Witziers & Van Vilsteren, 1990). Voor deze verschillen worden vaak de helderheid en de wijze waarop de leerstof binnen een vak is opgebouwd verantwoordelijk gesteld. Zo wordt het vak wiskunde - in tegenstelling tot het vakgebied geschiedenis - veelal gezien als een helder omschreven vak met een leerstofopbouw die sterk sequentieel en cumulatief van aard is en juist deze kenmerken worden als gunstig gezien voor de mogelijkheden om het onderwijs te coördineren. De positie van het vakgebied Engels is wat dit betreft wat onduidelijk, al zijn er aanwijzingen in de literatuur dat het leren en onderwijzen van vreemde talen op deze punten eerder vergelijkbaar is met wiskunde dan met geschiedenis (Els, Extra & Van Os, 1977). Daarnaast speelden de verwachte omvang van de vaksecties en de status van de vakgebieden binnen het schoolcurriculum een rol. Zo zijn secties Engels op dit punt interessanter dan secties Frans.

Een tweede relevante vraag in dit kader was welke docenten bij het onderzoek betrokken moesten worden. Hoewel het als wenselijk werd gezien dat alle docenten van een vaksectie aan het onderzoek meedoen, werd op basis van ander onderzoek (Hylkema, 1990) deze kans niet als groot ingeschat. Als zodanig heeft dit geleid tot het besluit om (bepaalde) secties de gelegenheid te geven alleen met een subgroep van docenten aan het onderzoek mee te doen in de hoop op deze manier de deelname van secties en de respons van docenten te optimaliseren. Uiteraard dient men dan deze groep wel nader te specificeren en in dit opzicht is de keuze gevallen op die docenten die (mede) werkzaam zijn op een HAVO-afdeling. Deze keuze is ingegeven door het feit dat bij het HAVO-onderwijs in principe de meeste docenten betrokken zijn. Immers, zowel eerstegraders als tweedegraders zijn bevoegd voor het lesgeven op deze afdelingen. Als vanzelfsprekend betekent deze keuze van docenten ook een beperking van de operationele populatie tot alle scholen met een HAVO-afdeling.

Dit laatste aspect (alle scholen met een HAVO-afdeling) duidt erop dat in principe naar omvang verschillende scholen tot de potentiële onderzoekspopulatie behoren. Dit is

niet geheel van belang ontbloot daar in de organisatieliteratuur de omvang van de organisatie altijd een belangrijke rol speelt bij het verklaren van verschillen in coördinatiestructuur. Eerder is dit naar voren gekomen bij de bespreking van de Marxse modellen. Daarnaast kan gewezen worden op auteurs als Mintzberg (1979) en Kieser en Kubicek (1983) die stellen dat, naarmate de organisatie groter is, de behoefte aan coördinatie toeneemt. Als zodanig wees dit gegeven op het belang de feitelijke steekproef zodanig vorm te geven dat secties en scholen van verschillende omvang bij het onderzoek betrokken zouden zijn.

3.4 Steekproeftrekking

De feitelijke steekproeftrekking bestond uit een tweetal stappen. Van belang hierbij is dat het onderzoek mede tot doel had op één vakgebied de relatie tussen coördinatie van het vakonderwijs en leerlingprestaties nader vast te stellen. Als zodanig veronderstelt een dergelijk doel dat valide en betrouwbare gegevens over achtergrondkenmerken van leerlingen en hun leerprestaties beschikbaar zijn. Om deze gegevens op efficiënte te verkrijgen is daarom aansluiting gezocht bij het onderzoek 'Uitgangssituatiemeting eerste fase voortgezet onderwijs' uitgevoerd door het CITO (Kremers, 1990). Dit onderzoek vond plaats in 1990 en had als doel het in kaart brengen van de leerresultaten aan het eind van het derde leerjaar in de vakgebieden biologie, Engels, Nederlands en wiskunde. Tevens zijn in het kader van dit onderzoek gegevens verzameld over achtergrondkenmerken van leerlingen.

Met betrekking tot de bovengenoemde vakgebieden werd, enigszins arbitrair, besloten om de relatie tussen coördinatie en leerlingprestaties nader te onderzoeken op het vakgebied wiskunde. Daartoe werden die vaksecties benaderd die meegedaan hadden aan bovengenoemd onderzoek. Van de 93 benaderde secties hebben in totaal 39 secties aan dit onderzoek meegedaan, waaronder 20 HAVO-secties en 19 MAVO-secties. Dit laatste wijst er op dat, vanwege de tegenvallende respons bij de HAVO-secties, bij deze groep afgeweken is van het oorspronkelijke plan om het onderzoek alleen te beperken tot docenten (mede) werkzaam op HAVO-afdelingen. Zoals bij de vraagstelling duidelijk is geworden, zal in deze

bijdrage aan de relatie tussen coördinatiemarken en leerlingprestaties verder geen aandacht geschonken worden.

De tweede stap bestond uit het werven van secties Engels en geschiedenis. Hiertoe is een bestand gevormd met alle scholen met een HAVO-afdeling. Vervolgens is deze groep gestratificeerd op grond van een tweetal variabelen, te weten schoolgrootte (drie strata) en regio (vier strata). De volgende stap bestond uit het berekenen van het aantal benodigde secties per cel, uitgaande van een gewenst aantal van 80 secties (gelijk verdeeld over de vakken Engels en geschiedenis). Ten slotte werd per cel het dubbele van het aantal benodigde secties at random benaderd. Gezien de tegenvallende respons in de eerste steekproeftrekking werd deze procedure herhaald, uiteraard met inachtneming van de resultaten van de eerste steekproeftrekking.

In totaal zijn voor het werven van vaksecties geschiedenis en Engels 198 scholen benaderd en daarinbinnen 350 secties. Hiervan deden in totaal 34 secties Engels en 31 secties geschiedenis, verspreid over 58 scholen, mee. Te zamen met de 39 vaksecties wiskunde gaat het in totaal om 103 secties, verspreid over 97 scholen. Wat betreft de verdeling over de stratificatievariabelen kan vermeld worden dat deze afwijkt van de oorspronkelijke planning. Ten koste van de secties in kleinere scholen (minder dan 850 leerlingen) zijn secties in grote scholen (meer dan 1154 leerlingen) oververtegenwoordigd. Hetzelfde geldt voor de scholen in de regio Zuid-West Nederland. Deze zijn oververtegenwoordigd ten koste van de scholen in Noord-West Nederland.

Wat betreft de respons van docenten moet in ogenschouw genomen worden dat bepaalde secties de gelegenheid geboden is slechts met een subgroep van docenten aan het onderzoek mee te doen (zie boven). Deze keuze heeft er concreet toe geleid dat bij de verspreiding van de vragenlijsten aan de verantwoordelijke contactpersoon (meestal de vaksectieleider) werd gevraagd aan alle docenten een vragenlijst te geven in het geval zijn sectie kleiner was dan zes personen en alleen aan die docenten die (mede) werkzaam waren op de HAVO- of MAVO-afdelingen wanneer dit niet het geval zou zijn. Deze werkwijze leidde er wel toe dat onbekend was hoeveel vragenlijsten precies

uitgezet zijn. Wel kon op basis van het aantal geretourneerde vragenlijsten ($n=376$) en gegevens over de omvang van de secties berekend worden dat 66% van alle bij deze secties werkzame docenten een vragenlijst ingevuld heeft.

3.5 Concepten en variabelen

Om de aard en mate van coördinatie van het vakonderwijs binnen scholen voor voortgezet onderwijs te kunnen onderzoeken is een analysekader nodig met behulp waarvan dit verschijnsel in kaart gebracht kan worden. Daartoe is aansluiting gezocht bij de literatuur over coördinatiemechanismen, waarbij vooral het werk van Mintzberg (1979, 1983) en van Kieser en Kubicek (1983) invloedrijk is geweest. Concreet heeft dit geleid tot een zestal concepten die elk verwijzen naar een specifieke wijze van coördinatie. Vervolgens is elk concept uiteengelegd in een aantal variabelen dat als indicator gezien wordt van het desbetreffende concept. Voor deze nadere invulling is vooral gebruik gemaakt van de meer specifiek – al dan niet expliciet – op coördinatie binnen scholen gerichte literatuur. In het navolgende zullen deze concepten en de (belangrijkste) variabelen die daaronder vallen kort beschreven worden.

Het eerste concept (*beleid*) heeft betrekking op het door vaksecties gevoerde onderwijsbeleid. In feite gaat het hier om de mate waarin secties expliciet vastleggen wat de aard en bedoeling is van hun onderwijs. Dit is onder andere vastgesteld door na te gaan welke (bindende) afspraken sectieleden maken ten aanzien van hun onderwijs en hoeveel gemeenschappelijke proefwerken door de sectie per leerjaar worden gegeven. Vanuit deze laatste variabele wordt duidelijk dat dit concept verwantschap vertoont met het in de coördinatie-literatuur gebruikte begrip standaardisatie (Mintzberg, 1979). Bij gemeenschappelijke proefwerken gaat het immers veelal om een standaardisatie van doelen, inhoud en toetsing van het onderwijs.

Het tweede concept (*overleg en samenwerking*) is gerelateerd aan het door Mintzberg (1979) gebedigde begrip wederzijdse afstemming; de afstemming van de activiteiten van de organisatieleden tijdens het werkproces door middel van onderling overleg en informatie-overdracht. Variabelen die hier een rol spelen

zijn onder andere de frequentie van overleg, de aard van overleg en het samenwerkingsklimaat. Deze laatste variabele geeft inzicht in de openheid van de communicatie, de kwaliteit van de onderlinge samenwerking en de mate waarin er binnen secties een zekere controle bestaat op elkaars werk.

Het concept *consensus* bouwt voort op de notie dat de mate waarin organisatieleden overeenkomstig denken en handelen een belangrijk vorm van coördinatie is. Mintzberg (1979) reserveert hiervoor het begrip coördinatie door middel van 'ideology'. Dit concept wordt geïndiceerd door een viertal variabelen die meten in hoeverre docenten binnen een sectie gemeenschappelijk denkbeelden koesteren over zaken als de taken en functie van de school, vakdoelen en -didactiek en de (gewenste) leerinhouden.

Het concept *besluitvorming* sluit aan bij de gedachte dat coördinatie nauw verbonden is met de besluitvormingsstructuur. Kieser en Kubicek (1983) stellen dat in veel organisaties leidinggevend verantwoordelijk zijn voor de coördinatie en daartoe uitgerust zijn met een bepaalde beslissingsbevoegdheid. Deze bevoegdheid kan echter ook toegewezen worden aan specifieke organen en/of individuen die formeel gezien geen hiërarchische functie vervullen. In veel schoolorganisatiemodellen komt dit laatste naar voren in de nadruk die gelegd wordt op de professionele autonomie van docenten (Mintzberg, 1979). Andere auteurs wijzen er daarentegen op dat ook andere participanten in de besluitvorming – zoals vaksecties – een belangrijke invloed kunnen hebben op de dagelijkse gang van zaken in het onderwijs (Marx, 1975). In dit opzicht mag dus verwacht worden dat scholen op dit punt zullen verschillen. Het concept besluitvorming is uiteengelegd in een tweetal variabelen die inzicht geven in de mate waarin individuele docenten zelfstandig kunnen beslissen over hun onderwijs – de befaamde professionele autonomie – dan wel de mate waarin de vaksectie hier een belangrijke rol speelt.

De laatste twee concepten (*schoolleiding en sectieleider*) hebben respectievelijk betrekking op de onderwijskundige rol van de schoolleiding en van de vaksectieleider. Deze concepten bouwen voort op begrippen als 'direct supervision' (Mintzberg, 1979) en 'Koördination

durch persoonlijke Weisungen' (Kieser & Kubicek, 1983). Voor een nadere invulling van het eerste concept is uitgegaan van het begrip onderwijskundig leiderschap zoals dat door Firestone en Wilson (1985) is uitgewerkt. Zij onderscheiden een drietal wijzen waarop de schoolleiding het onderwijs kan coördineren, waaronder controle en regelgeving. Voor het tweede concept is aansluiting gezocht bij Smets (1986) die verschillende rollen onderscheidt die een vaksectieleider kan vervullen. Hier gaat het om de mate waarin de sectieleider zich als onderwijskundig leider van de sectie manifesteert. Bij beide concepten spelen daarbij nog variabelen een rol die inzicht geven in de mate van invloed van deze actoren op een aantal vakinhoudelijke beslispunten.

Een overzicht van bovenstaande concepten

en variabelen wordt weergegeven in Tabel 1. Tevens geeft deze tabel inzicht in het aantal per variabele ontwikkelde items, de betrouwbaarheid van de afzonderlijke schalen, de gemiddelde scores, de schaalbreedtes, de ladingen van de variabelen op de concepten en bij wie welke informatie verzameld is. Op deze twee laatste aspecten zal in volgende paragrafen nog teruggekomen worden.

Voor nadere informatie over de inhoud van de vragen wordt verwezen naar het oorspronkelijke onderzoeksrapport (Witziers, 1992). Een uitzondering wordt hier gemaakt voor een aantal schalen dat betrekking heeft op het begrip consensus. Drie van de vier schalen zijn op vakspecifieke wijze geoperationaliseerd. Met andere woorden, voor elk vakgebied zijn aparte schalen geconstrueerd. In dit opzicht zijn de ge-

Tabel 1
Overzicht concepten en variabelen

Concept (variabelen)	aantal items	α	gemiddelde score	schaalbreedte	parameter waarde	info bron
Beleid						
sectieregels	17	0.84	2.60	1-4	0.82	VL
gemeenschappelijke proefwerken	4	0.94	5.24*	nvt	0.54	VL
planning	8	0.91	2.23	1-4	0.33	VL
schoolwerkplan	3	0.73	1.31	1-3	0.52	VL
Overleg en samenwerking						
frequentie overleg	1	nvt	18.92**	nvt	0.38	VL
aard formeel overleg	14	0.83	2.57	1-4	0.50	VL
aard informeel overleg	14	0.79	2.60	1-4	0.59	D
cohesie	11	0.85	2.98	1-4	0.20	D
samenwerkingsklimaat	11	0.82	2.72	1-4	0.87	D
Consensus						
homogeniteit taken/functie school	14	0.74	0.27	0-2.31***	0.73	A
homogeniteit doelen vakonderwijs	10	0.81	1.41	0-5.77***	0.92	A
homogeniteit in lesgedrag	12	0.85	1.55	0-5.77***	0.85	A
homogeniteit leerinhouden	5	0.77	5.27	nvt	0.73	A
Besluitvorming						
invloed docent	14	0.86	2.91	1-4	-0.40	A
invloed vaksectie	14	0.86	2.77	1-4	0.96	A
Schoolleiding						
controle schoolleiding	17	0.92	1.43	1-4	0.88	VL
verticale communicatie	17	0.91	1.61	1-4	0.79	VL
schoolregels	17	0.90	1.88	1-4	0.63	VL
invloed schoolleiding	14	0.78	1.40	1-4	0.52	VL
Vaksectieleider						
rol vaksectieleider	3	0.79	2.13	1-4	0.66	D
invloed vaksectieleider	14	0.94	1.51	1-4	0.63	A

* = aantal gemiddelde proefwerken per leerjaar

** = aantal uren overleg op jaarbasis

*** = gebaseerd op minimale-maximale spreiding

VL = informatie ingewonnen bij vaksectieleider

D = informatie ingewonnen bij docenten

A = informatie ingewonnen bij alle betrokkenen

gevens over deze schalen in Tabel 1 in zekere zin misleidend, daar het hier alleen gaat om gegevens die betrekking hebben op de schalen ontwikkeld voor het vakgebied wiskunde.

3.6 Dataverzameling en -verwerking

Bij de dataverzameling heeft niet elke respondent met dezelfde vragenlijst te maken gehad. Tabel 1 (laatste kolom) geeft daar nader inzicht in. De reden hiervoor was dat enerzijds zoveel mogelijk informatie ingewonnen moest worden, anderzijds moest dit niet leiden tot uitzonderlijk lange vragenlijsten. Dit laatste zou immers de kans op non-respons alleen maar vergroten. Bij de verwerking van de gegevens zijn de gegevens die afkomstig waren van meerdere respondenten binnen een school geaggregeerd. Veelal betekent dit dat de gemiddelde score van deze docenten gebruikt is als score van een sectie. Een uitzondering hierop vormen de schalen die betrekking hebben op het concept consensus. Daarbij was namelijk van belang dat docenten konden aangeven in welke mate zij het eens waren met uitspraken die betrekking hadden op de doelen van het (vak)onderwijs, hun wijze van lesgeven en de gewenste (vak)inhouden. Als sectiemaat is hier gebruik gemaakt van de gemiddelde spreiding in de antwoorden van docenten per item. Als zodanig betekent het gebruik van de spreiding tussen docenten als maat dat een lage score een hoge mate van consensus betekent.

3.7 Analysemethoden

Zoals uit het bovenstaande naar voren is gekomen zijn de data per variabele geaggregeerd tot school/sectieniveau. Vervolgens zijn over deze scores de gemiddelde scores berekend. Deze gemiddelden hebben vooral een rol bij de beantwoording van de vraag naar de aard en mate van coördinatie van het vakonderwijs. Aangezien bij de beantwoording van een descriptieve vraag vaak ook gekeken wordt of subpopulaties verschillen op de onderzochte kenmerken, is verder nagegaan in hoeverre er tussen scholen en vaksecties op dit punt verschillen waren te constateren. Concreet betekent dit dat met behulp van een eenwagsvariantie-analyse onderzocht is of de scores naar vakgebied en naar school- en sectieomvang significant van elkaar verschilden. De omvang van scholen is geoperationaliseerd

door middel van het aantal leerlingen, de omvang van de vaksectie door middel van het aantal leden.

Om te komen tot een typologie van scholen - de tweede vraag - is gebruik gemaakt van een clusteranalyse. Een dergelijke techniek is geschikt om te onderzoeken of er in de empirie op basis van onderzochte kenmerken duidelijke herkenbare typen terug te vinden zijn. Een probleem bij deze analyse was het grote aantal variabelen dat oorspronkelijk onderscheiden is; het zoeken naar typen is gebaat bij een relatief klein aantal kenmerken (Everitt, 1974). Voor deze clusteranalyse is dan ook geen gebruik gemaakt van de afzonderlijke variabelen, maar van de onderscheiden concepten. Uit Tabel 1 (vijfde kolom) kan afgeleid worden hoe de score van een sectie op deze concepten bepaald is. Zo is een score van een sectie op het concept beleid gelijk aan de som van de gewogen scores op de indicatoren sectieregels, gemeenschappelijke proefwerken, planning en schoolwerkplan. Deze gewogen scores zijn verkregen door middel van een LISREL-analyse waarover hier niet nader wordt gerapporteerd.

Van belang om te vermelden is ten slotte dat voor elke groep van secties afzonderlijk een clusteranalyse is uitgevoerd. Gezien het feit dat het gaat om secties op een drietal vakgebieden, betekent dit concreet dat een drietal clusteranalyses is uitgevoerd. Een belangrijke reden hiervoor was het feit dat de verschillende schalen die betrekking hebben op het concept consensus op vakspecifieke wijze zijn geoperationaliseerd. Dit betekent dat de schalen niet per se onderling vergelijkbaar zijn. Een en ander impliceerde, dat wou dit concept betrokken worden bij een clusteranalyse, een drietal verschillende analyses noodzakelijk was. Een nadeel van deze werkwijze was wel dat het steeds ging om analyses toegepast op een relatief klein aantal secties. Als voordeel kon echter gezien worden dat, door het toepassen van een drietal analyses, er een vorm van kruisvalidatie mogelijk werd; de zeggingskracht van de resultaten is groter, naarmate de verschillende analyses op de verschillende vakgebieden verwante resultaten laten zien.

4 Resultaten

4.1 Beschrijvende gegevens

De eerste vraagstelling heeft – zoals gezegd – betrekking op de aard en mate van coördinatie van het vakonderwijs. In de inleiding is daarbij naar voren gekomen dat op basis van de dominante visie op het functioneren van schoolorganisaties verwacht mocht worden dat coördinatie van het vakonderwijs geen grote rol speelt binnen scholen van voortgezet onderwijs. In operationele zin luidde dan ook de verwachting, dat de scores op alle onderscheiden variabelen laag zouden zijn, uitgezonderd de variabele invloed van de individuele docent en de variabelen die betrekking hebben op het concept consensus. Gezien het accent dat in de literatuur gelegd wordt op de autonomie van docenten en de verschillen in visie op het onderwijs werden hier juist hoge scores verwacht. Voor de hand ligt hier de opmerking dat het bij de begrippen 'hoog' en 'laag', vanwege het ontbreken van absolute maatstaven, met name gaat om een interpretatie van de kant van de onderzoeker.

De resultaten van het beschrijvend gedeelte van het onderzoek kunnen gedeeltelijk ontleend worden aan de gemiddelde scores op de verschillende schalen (zie Tabel 1). Een belangrijk resultaat hier is dat de homogeniteit in opvattingen tussen docenten aanzienlijk is. Zo is de gemiddelde spreiding tussen docenten per sectie op de veertien items die de mate van homogeniteit in opvattingen over de taken en de functie van de school representeren slechts

0.27. Ook uit de resultaten aangaande de andere variabelen die op dit concept betrekking hebben komt dit beeld naar voren. Als zodanig geven deze resultaten aanleiding om te veronderstellen dat verschillen in opvattingen en gedrag binnen een sectie niet erg groot zijn. Andere belangrijke resultaten die ontleend kunnen worden aan Tabel 1 zijn daarnaast dat vaksecties regelmatig overleggen (circa 19 uur op jaarbasis), dat secties per leerjaar gemiddeld vijf gemeenschappelijke proefwerken geven en dat zowel schoolleiding als vaksectieleider weinig coördinerende taken ter hand nemen. Op basis van deze gegevens zou men kunnen concluderen dat vaksecties organen zijn die hun onderwijs (relatief) veel coördineren, terwijl de rol van zowel schoolleiding als vaksectieleider in deze niet erg groot is.

Een genuanceerder overzicht aangaande coördinatie van het vakonderwijs ontstaat echter als de data aan een meer nauwgezette blik worden onderworpen. Een voorbeeld op dit punt zijn bijvoorbeeld de data die inzicht geven in de mate van invloed van verschillende actoren binnen de schoolorganisatie op een aantal belangrijke vakinhoudelijke beslispunten. In Tabel 2 worden deze data weergegeven.

Uit deze Tabel blijkt dat vaksecties vooral belangrijke organen zijn bij de keuze van de leergang, voor het vastleggen van de leerdoelen en de inhoud van het onderwijs, en voor de toetsing. Verder blijkt dat de invloed van de schoolleiding en vaksectieleider op de meeste punten te verwaarlozen is en dat de invloed van de individuele docent vooral tot uiting komt op

Tabel 2

Door docenten gepercipieerde invloed van verschillende factoren binnen de schoolorganisatie op het onderwijs; gemiddelden; n=367

	school leiding	vaksectie- leider	vaksectie	docent	overige personen en organen
uitgangspunten onderwijs	2.50	1.54	2.33	2.20	1.86
keuze leergang	1.11	1.73	3.82	2.08	1.04
vaststellen leerdoelen	1.40	1.63	3.24	2.62	1.27
leerstofkeuze	1.07	1.63	3.45	2.66	1.05
wijze van lesgeven	1.23	1.42	2.34	3.57	1.11
pedagogische aanpak	1.48	1.33	1.87	3.64	1.27
inhoud proefwerken	1.01	1.36	2.92	3.21	1.04
aantal proefwerken	1.65	1.39	2.80	3.03	1.17
normering proefwerken	1.04	1.32	2.92	3.12	1.03
rapportcijfer	1.46	1.40	2.83	3.16	1.27
inhoud schoolonderzoek	1.29	1.44	3.15	2.64	1.04
hoeveelheid huiswerk	1.22	1.17	1.74	3.80	1.12
vergaderfrequentie	1.71	2.48	2.93	1.76	1.15

Schaalbreedte loopt van 1 (niet of nauwelijks van invloed) tot 4 (zeer grote invloed)

het pedagogisch-didactisch terrein. Verder blijkt dat de invloed van andere personen en organen binnen de school op het (vak)onderwijs zeer gering is.

Vooruitlopend op de conclusies betekenen deze resultaten een tweetal zaken. Enerzijds dat docenten weliswaar een zekere autonomie hebben, maar dat deze gestalte krijgt binnen een door de vaksectie vastgesteld kader. Anderzijds dat de coördinerende rol van leidinggevenden wat betreft het vakonderwijs als gering ingeschat moet worden, al kan dit genuanceerd worden door te wijzen op de invloed die de schoolleiding heeft op de uitgangspunten van het onderwijs.

Deze conclusies worden ondersteund als gekeken wordt naar de overige resultaten van dit onderzoek. Uit de gegevens over regelgeving op sectie- en schoolniveau (zie Tabel 3) blijkt dat vaksecties het gedrag van docenten sterk reguleren op punten als de leerdoelen, het gebruik van de leergang, de inhoud van het onderwijs en de wijze en aard van toetsing, terwijl de pedagogisch-didactische aanpak van docenten veel minder aan regelgeving gebonden is. Een verwant patroon komt naar voren bij regelgeving op schoolniveau, zij het dat deze scores in het algemeen veel lager zijn, uitgezonderd de mate van regelgeving op schoolniveau ten

aanzien van de uitgangspunten van het onderwijs.

Verdere ondersteuning voor bovengenoemde conclusies kan ten slotte ontleend worden aan de gegevens over de aard van het (in)formele overleg tussen docenten. Hieruit blijkt dat veel overlegd wordt over aspecten als de planning van het onderwijs, de realisatie hiervan, de wijze en inhoud van toetsing, het tempo van lesgeven en de inhoud van het onderwijs, terwijl weinig overleg plaatsvindt over aspecten als het didactisch functioneren van docenten en de problemen die docenten in de dagelijkse onderwijspraktijk ondervinden.

4.2 Verschillen tussen scholen en secties

Een ander belangrijk aspect in dit kader is de vraag of er naar aard en mate van coördinatie van het vakonderwijs tussen scholen en secties verschillen te onderkennen zijn. Zoals gezegd, is om deze vraag te onderzoeken een eenwegsvariantie-analyse toegepast, waarbij als factoren gediend hebben het vakgebied waarop vaksecties opereren, de omvang van de sectie (aantal docenten) en de omvang van de school (aantal leerlingen). Een gedeelte van deze resultaten staat vermeld in Tabel 4.

Uit deze tabel blijkt dat er inderdaad belangrijke verschillen in functioneren tussen vaksecties te onderkennen zijn. Met name het vakgebied waarop de sectie opereert is in dit kader van belang. Secties Engels komen naar voren als die vaksecties die het meest hun onderwijs coördineren. De rol van de vaksectieleider is bij deze secties geprononceerder dan bij de overige secties, zij kennen een grotere invloed van de vaksectie op de onderwijsbesluitvorming en zij reguleren sterker elkaars handelen, met name met behulp van gemeenschappelijke proefwerken. Daarbij kan opgemerkt worden dat deze secties op deze punten veel meer verschillen van vaksecties geschiedenis dan van vaksecties wiskunde. In dit opzicht kan geconcludeerd worden dat deze resultaten in zekere zin afwijken van de eerder geformuleerde verwachtingen omtrent de verschillen tussen vakken. Conform de verwachting is de relatief zwakke rol van vaksecties geschiedenis, veel minder het gegeven dat vaksecties Engels niet alleen secties geschiedenis overtreffen, maar ook vaksecties wiskunde.

Tabel 3
Regelgeving binnen scholen; gemiddelden; $n=103$

	sectie-niveau	school-niveau
uitgangspunten onderwijs	2.37	2.51
leerdoelen	3.32	2.29
leergang	3.83	2.22
leerstofkeuze	3.67	2.22
volgorde onderwerpen	2.81	1.71
concrete leerstof lessen	2.17	1.39
opgaven/sommen in lessen	2.11	1.41
tijd te besteden aan onderwerpen	2.43	1.57
tot stand koming rapportcijfer	3.30	2.64
normering proefwerken	2.88	1.79
inhoud proefwerken	2.70	1.56
frequentie proefwerken	2.81	2.48
frequentie overhoringen	1.88	1.58
didactische werkwijzen	2.05	1.66
hoeveelheid huiswerk	1.67	1.65
na- en bijscholing leraren	1.36	1.48
vergaderfrequentie	2.75	2.10

Schaalbreedte loopt van 1 (leraar handelt volledig op basis van eigen inzichten) tot 4 (handelen van leraar in hoge mate bepaald door afspraken)

Tabel 4

Verschillen in coördinatie naar vakgebied en omvang vaksectie: eenwegvariantie-analyse, Scheffe ($\alpha=.05$); $n=103$

Concept (variabelen)	Engels	wiskunde	geschiedenis	omvang (<4)	omvang (4-6)	omvang (>6)
Beleid						
sectieregels	2.95	2.53	2.45	2.48	2.54	2.82
gemeenschappelijke proefwerken	6.05	6.02	3.66*	6.30	4.68	5.18
planning	2.26	2.05	2.41	2.31	2.24	2.12
schoolwerkplan	1.37	1.23	1.34	1.33	1.26	1.36
Overleg en samenwerking						
frequentie overleg	24.28	14.83	17.96*	12.31	18.77	22.92*
aard formeel overleg	2.60	2.54	2.58	2.52	2.55	2.63
aard informeel overleg	2.67	2.64	2.49*	2.49	2.67	2.64
cohesie	2.92	2.97	3.00	3.16	2.96	2.87
samenwerkingsklimaat	2.80	2.73	2.64	2.74	2.69	2.75
Consensus						
homogeniteit taken/functie school	0.27	0.27	0.27	0.26	1.27	0.28
homogeniteit doelen vakonderwijs	1.47	1.37	1.40	1.37	1.37	1.48
homogeniteit in lesgedrag	1.70	1.57	1.39*	1.45	1.50	5.39
homogeniteit leerinhouden	6.40	4.87	6.10*	5.57	5.39	6.18
Besluitvorming						
invloed docent	2.67	2.88	2.97*	2.93	2.99	2.76
invloed vaksectie	3.09	2.67	2.58	2.75	2.72	2.90
Schoolleiding						
controle schoolleiding	1.29	1.43	1.56	1.41	1.40	1.30
verticale communicatie	1.48	1.67	1.69	1.65	1.60	1.59
schoolregels	1.78	1.98	1.98	1.90	1.78	2.10
invloed schoolleiding	1.40	1.40	1.38	1.42	1.38	1.41
Vaksectieleider						
rol vaksectieleider	2.23	2.15	1.99*	2.14	2.12	2.13
invloed vaksectieleider	1.55	1.54	1.43	1.60	1.47	1.49

* significante verschillen tussen secties

Met betrekking tot omvang van de school en de vaksectie kan ten slotte gesteld worden dat deze variabelen geen grote rol spelen bij het verklaren van verschillen in aard en mate van coördinatie. Wat betreft de omvang van de vaksectie (zie Tabel 4) is slechts bij één variabele, te weten de frequentie van het formele overleg, sprake van een significant verschil: docenten van grote secties overleggen (veel) meer dan docenten van kleine secties. Bij de variabele omvang van de school (niet in tabelvorm weergegeven) is geen enkel verschil tussen scholen van verschillende omvang significant. Als zodanig geven deze gegevens aanleiding tot de conclusie dat de omvang van de organisatie, anders dan in de literatuur wordt verondersteld, nauwelijks een relatie vertoont met de coördinatiestructuur.

4.3 Een typologie van scholen

De tweede vraagstelling van dit onderzoek luidde of er naar coördinatie van het vakonder-

wijs in de empirie verschillende typen scholen te onderkennen zijn die overeenkomen met de Marxse modellen. Deze vraag is onderzocht met behulp van clusteranalyses die op elk vakgebied zijn toegepast. Specifieker gaat het om een clustermethode Ward die gebruik maakt van de gekwadraterde euclidische afstandsmaat.

In Tabel 5 staan de resultaten van de clusteranalyse toegepast op de wiskundesectionen. Uit deze tabel blijkt dat gekozen is voor een drie clusteroplossing. Deze oplossing voldoet goed aan criteria als een evenredige verdeling van de onderzoekseenheden over de verschillende clusters en homogeniteit (Everitt, 1984). Tevens komt in deze tabel naar voren welke variabelen significant van nul (de gemiddelde score) afwijken en waar er significante verschillen tussen de clusters bestaan. Voor de interpretatie van Tabel 5 is verder van belang te vermelden dat de scores op de consensusschaal zodanig zijn bewerkt dat een hoge positieve waarde

op dit kenmerk een grote mate van overeenstemming impliceert.

Tabel 5
Resultaten clusteranalyse, vakgebied wiskunde, methode Ward (3 clusters, gemiddelde scores op de kenmerken binnen clusters, significante verschillen tussen clusters); $n=39$

	cluster 1 ($n=14$)	cluster 2 ($n=15$)	cluster 3 ($n=10$)
Beleid	-.91	.72	.18**
Overleg	-.81	.43	.47*
Consensus	-.19	.39	-.29
Besluitvorming	-.77	.26	.63*
Schoolleiding	-.65	.89	-.44*
Sectiehoofd	-.21	.83	-.95**

* significante verschillen tussen cluster 1 en 3 en cluster 2 en 3

** significante verschillen ($p < 0.05$) tussen elke combinatie van clusters

vetgedrukt: scores die significant ($p < 0.10$) afwijken van nul

Uit Tabel 5 blijkt dat er inderdaad sprake is van inhoudelijk duidelijk te onderscheiden clusters. Met name geldt dit voor de eerste twee clusters, waarbij het eerste cluster overwegend lage scores kent, het tweede cluster overwegend hoge scores. Meer specifiek, het eerste cluster wordt gekenmerkt door vaksecties met weinig onderwijsinhoudelijk overleg en onderwijskundig beleid, terwijl docenten beslissingen aangaande het onderwijs op betrekkelijk autonome wijze kunnen nemen. Verder wordt dit cluster gekenmerkt door een zwakke coördinerende rol van de schoolleiding.

Het tweede cluster vertoont daarentegen (grotendeels) tegenovergestelde kenmerken. Het derde cluster neemt tegen deze achtergrond een middenpositie in; enerzijds is de coördinerende rol van het (midden)management zwak ontwikkeld, anderzijds overleggen secties veel en heeft de vaksectie een grote invloed op beslissingen aangaande het onderwijs. Als zodanig vormen ons inziens deze resultaten een gedeeltelijke bevestiging van de empirische validiteit van de Marxse modellen. Immers, in het segmentale model zoals door diverse auteurs beschreven, komt naar voren dat segmentale scholen gekenmerkt worden door autonomie, gebrek aan overleg en samenwerking, gebrek aan consensus en zwak ontwikkeld leiderschap en juist deze kenmerken zijn typerend voor scholen in het eerste cluster. Anderzijds is er een groep van scholen (cluster 2)

die wat betreft deze kenmerken juist sterk afwijkt van het segmentale model en in dit opzicht ligt de overeenkomst met het collegiale en/of coöperatief regulerend model voor de hand. Problematischer is uiteraard wel het derde cluster. Dit cluster is moeilijk in Marxse begrippen te vangen.

Een belangrijke vervolgvraag is dan of de clusteranalyses toegepast op de andere vaksecties (Engels en geschiedenis) een vergelijkbaar patroon laten zien. Deze vraag kan in grote lijnen positief beantwoord worden. Uitgaande van een drie clusteroplossing, zijn op beide vakgebieden een tweetal clusters te onderkennen die verwantschap vertonen met de bovenbeschreven clusters een en twee. Anders gesteld, enerzijds is er op elk vakgebied een cluster te onderkennen met voor het merendeel scores ver beneden het gemiddelde, anderzijds een cluster met grotendeels scores ver boven het gemiddelde. Daarnaast is er op deze beide vakgebieden steeds een derde cluster aanwezig, dat tussen deze twee clusters een middenpositie lijkt in te nemen. Ten slotte verschillen ook op deze vakgebieden de clusters significant van elkaar op het merendeel van de onderzochte kenmerken. In dit opzicht komen deze resultaten dus in grote lijnen overeen met de resultaten van de vaksecties wiskunde.

5 Conclusies en discussie

In dit artikel stond een tweetal vragen centraal. Bij de eerste vraag ging het om de aard en mate van coördinatie van het vakonderwijs binnen scholen voor voortgezet onderwijs. Het antwoord op deze vraag werd vooral gezocht in een beschrijving van de wijze waarop vaksecties en (midden) management in scholen voor voortgezet onderwijs het vakonderwijs coördineren. De tweede vraag had betrekking op de vraag of er naar coördinatie verschillende typen scholen te onderkennen zijn die in meer of mindere mate overeenkomen met de Marxse modellen.

Ten aanzien van de eerste vraag is vooraf op basis van modellen als de school als losjes gekoppeld systeem en de school als professionele bureaucratie de verwachting geformuleerd dat zowel vaksecties als (midden)management een geringe rol zouden spelen bij de coördinatie

van het vakonderwijs. Gegeven de gepresenteerde resultaten ligt de conclusie voor de hand dat deze stelling vooral ten aanzien van de wijze waarop vaksecties functioneren onjuist is. In tegenstelling tot de geformuleerde verwachtingen wordt er binnen vaksecties veel overlegd, worden er veel afspraken betreffende het onderwijs gemaakt worden, is er binnen secties sprake van relatief homogene onderwijsopvattingen en heeft de vaksectie een grote invloed op onderwijskundige beslissingen. Meer in lijn met bovengenoemde modellen is het gegeven dat de coördinerende rol aangaande het vakonderwijs van zowel schoolleiding als vaksectieleiders van ondergeschikt belang is. Dit blijkt bijvoorbeeld uit de geringe invloed die zowel schoolleiding als vaksectieleider hebben op de aard, doelen en inhoud van het onderwijs. Daarnaast is er op schoolniveau betrekkelijk weinig regelgeving om het gedrag van docenten en secties te reguleren en zijn de contacten tussen schoolleiding en docenten weinig intensief.

Deze meer algemene conclusie kan genuanceerd worden door te wijzen op de gerichtheid van de coördinatie. Vrijwel alle resultaten wijzen op het feit dat coördinatie vooral gestalte krijgt in de vorm van standaardisatie van de inhoud en de uitkomsten van het werk. Bij vaksecties blijkt dit gegeven bijvoorbeeld uit het feit dat kenmerken als regelgeving, overleg en besluitvorming zich centreren rondom aspecten als het vastleggen van de leerdoelen en de leerstof en de toetsing daarvan. Een tweede nuancering vloeit voort uit het feit dat er tussen vakken verschillen zijn te onderkennen. Vaksecties Engels coördineren hun onderwijs meer dan vaksecties wiskunde en, vooral, geschiedenis. In dit opzicht is er dus aanleiding om te veronderstellen dat de aard van het vak een rol van betekenis speelt bij het verklaren van verschillen in functioneren van vaksecties (Murphy, 1988; Hylkema, 1990). Aan de andere kant is het echter de vraag of dit een goede verklaringsgrond is. Ten eerste valt te wijzen op het feit dat er weliswaar verschillen zijn tussen vaksecties in verschillende vakgebieden, maar de vraag is of veel van die verschillen wettelijk zijn. Significante verschillen betekenen niet altijd grote verschillen. In de tweede plaats kan gewezen worden op de variatie in functioneren binnen een vakgebied. Met name de re-

sultaten van de clusteranalyses laten zien dat het type school waarbinnen de vaksectie functioneert belangrijk is. In dit opzicht ligt het voor de hand eerder verschillen tussen scholen dan tussen vakgebieden als verklaringsgrond te nemen voor verschillen in coördinatie van het vakonderwijs. Een derde nuancering ten slotte heeft betrekking op de rol van de schoolleiding. In het bovenstaande is gesteld dat de coördinerende rol van de schoolleiding ten opzichte van het vakonderwijs niet uit de verf komt. Een belangrijke nuancering is echter wel dat de invloed van de schoolleiding op de uitgangspunten van het door vaksecties gegeven onderwijs redelijk groot is. In de rede ligt dan ook de conclusie dat ook vaksecties werken binnen op schoolniveau vastgestelde kaders, zij het dat deze kaders waarschijnlijk veel minder strak zijn dan bijvoorbeeld een praktijk als het geven van gemeenschappelijke proefwerken. Niettemin is de vraag of een uitspraak van Weick (1984, p. 397) als zou er binnen scholen sprake zijn van een 'disconnection between headquarters and the field' niet al te sterk is.

Als zodanig geven deze resultaten ook aanleiding tot andersoortige beschouwingen. Elders (Witziers, 1992, 1993) is betoogd dat, gezien de belangrijke rol die vaksecties spelen bij de vormgeving en uitvoering van het onderwijs, variabelen die betrekking hebben op het functioneren van vaksecties in onderzoeken naar schooleffectiviteit opgenomen dienen te worden. Gezien de aard van deze bijdrage zal hier niet nader op worden ingegaan. Wel zal ingegaan worden op een andere consequentie en dat is dat het klassieke beeld dat binnen scholen docenten op individuele wijze hun onderwijs vormgeven onjuist is. Docenten in het voorgezet onderwijs vormen steeds meer een team dat als collectief beslissingen aangaande het onderwijs neemt. In algemene zin kan deze ontwikkeling gezien worden als een verschuiving in de machtsverhoudingen binnen scholen. Meer specifiek betekent dit een tweetal zaken. In de eerste plaats dat de docent nog slechts op een aantal punten volledige beslissingsbevoegdheid wordt gegund. Voor zover van dit laatste nog wel sprake is, moet deze beslissingsbevoegdheid vooral gezocht worden binnen het instructiedomein van de school. Op andere punten daarentegen moet de docent deze beslissingsbevoegdheid delen met anderen.

Gezien het feit dat binnen scholen voor voortgezet onderwijs coördinatie vooral betrekking heeft op leerdoelen, leerstof en de toetsing hiervan, kan opgemaakt worden dat het hier vooral gaat om beslissingen aangaande de kwalificatie en selectie van leerlingen. Verder betekent deze verschuiving dat het handelen van leerkrachten meer en meer onderworpen is aan het toezicht van collega's. Dit punt kan verduidelijkt worden aan de hand van de veelvoorkomende praktijk van het geven van gemeenschappelijke proefwerken. Verwacht mag worden dat in deze praktijk (mede) door leerkrachten zelf ontwikkelde normen aangaande 'goed' onderwijs tot uiting komen en zodoende een vorm van controle op elkaars werk inhouden. Voor de hand ligt dan de opmerking dat deze verschuiving in beschouwingen over de aard en het functioneren van schoolorganisaties meegenomen dient te worden.

Naast een theoretisch belang kan er aan de gevonden resultaten zeker ook een praktisch belang worden gehecht. Zo mag men verwachten dat veel vaksecties qua interne organisatie voldoende toegerust zijn om bijvoorbeeld de innovaties in het kader van de basisvorming in te voeren. Als condities als een goede overlegstructuur, consensus en een democratische besluitvorming van belang zijn voor de invoering van innovaties (Fullan, 1991), dan hoeft wat betreft deze aspecten de zaak niet somber ingeschat worden, zeker wat betreft de vakgerichte lijn. Voor zover er wel sprake is van een problematisch aspect, is dit het 'gesloten' klimaat binnen scholen/secties. Zo komen problemen die docenten ervaren binnen hun klassesituatie en het functioneren van docenten voor de klas in overlegsituaties nauwelijks aan bod. Hier liggen dan ook aanknopingspunten om het functioneren van scholen te verbeteren, al is duidelijk dat dit, gezien de hardnekkigheid waarmee dit probleem in de literatuur naar voren komt, geen eenvoudige zaak zal zijn.

Wat betreft de tweede vraag, die naar de verschillen tussen scholen, zijn de resultaten relatief eenduidig. De resultaten van de clusteranalyses laten zien dat er aan de ene kant van het spectrum scholen zijn waar het coördinerend vermogen van zowel vaksecties als (mid-den)management zwak ontwikkeld is, terwijl aan de andere kant van het spectrum scholen te onderscheiden zijn met tegenovergestelde ken-

merken. In dit opzicht passen deze resultaten dus goed bij het door Marx ontwikkelde gedachtegoed en is er een basis om op de onderzochte aspecten scholen te typeren in termen van segmentaal en, daartegenover, van coöperatief-gereguleerd. Uiteraard houdt het gebruik van deze begrippen niet de suggestie in dat de modellen waar deze begrippen betrekking op hebben in hun geheel getoetst zijn. Modellen van schoolorganisatie impliceren immers meer dan in het kader van dit onderzoek onderzocht is. Een belangrijk punt is verder dat in dit onderzoek per school maar steeds een sectie onderzocht is, waardoor nauwelijks een uitspraak gedaan kan worden over de vraag of binnen een bepaalde groep van scholen alle secties beter of slechter functioneren dan binnen een andere groep. De vraag of *scholen* dus verschillen naar coördinerend vermogen is in dit onderzoek dus eigenlijk niet beantwoord. Overigens is deze vraag wel beantwoord in het vooronderzoek. Uit dit case-study bleek dat scholen, ondanks verschillen in functioneren tussen secties binnen een school, systematisch verschillen in de wijze waarop secties hun onderwijs coördineren (Witziers & Van Vilsteren, 1990). Voor de hand ligt de opmerking dat deze resultaten vooral van nut kunnen zijn bij de begeleiding en advisering van scholen, te meer daar er aanwijzingen zijn dat scholen die afwijken van het segmentale model in termen van leerlingprestaties effectiever zijn (cf. Witziers, 1992, 1993).

6 Kanttekeningen

Een belangrijke vraag in het kader van dit artikel is of, en in hoeverre, de gevonden resultaten een waarheidsgetrouw beeld geven van de werkelijke situatie binnen scholen. Het belangrijkste probleem in dit opzicht vloeit ons inziens voort uit het feit dat de respons rate van het onderzoek laag is en uit het feit dat kleine scholen ondervetegenwoordigd zijn in de onderzoeks-populatie. Dit kan betekenen dat alleen die vaksecties aan het onderzoek meegedaan hebben die reeds tot een zekere wasdom gekomen zijn. Mocht dit het geval zijn, dan is duidelijk dat de onderzoeksresultaten vertekend zijn. Daar staat echter tegenover dat de resultaten van dit onderzoek in grote lijnen consistent zijn met

ander onderzoek op dit terrein (Friebel, 1994; Hylkema, 1990; Kuhlemeier & Van den Bergh, 1995; Siskin, 1994). Daarnaast valt te wijzen op de vele overeenkomsten tussen de hier gevonden resultaten en de resultaten uit het vooronderzoek (Witziers & Van Vilsteren, 1990; Van Vilsteren & Witziers, 1990a/b). In dit opzicht ligt wel degelijk de conclusie voor de hand dat de gevonden resultaten betreffende het functioneren van scholen en secties als valide te karakteriseren zijn.

Een tweede belangrijk probleem in dit onderzoek is dat per sectie een relatief klein aantal docenten betrokken is geweest bij het vaststellen van de organisatiekenmerken. In algemene zin betekent dit dat de betrouwbaarheid in het geding is, daar het vaststellen van organisatiekenmerken gebaat is bij zoveel mogelijk respondenten (Rowan, Raudenbusch & Kang, 1989). Het meest relevant is dit wellicht voor de schalen die betrekking hebben op het concept consensus. Dat door het geringe aantal deelnemende docenten een vertekening kan zijn opgetreden, is duidelijk, temeer daar de betrokken docenten naar het type onderwijsafdeling waar zij bij betrokken zijn ('MAVO-' versus 'HAVO-docenten') een vrij homogene groep vormen. Dit kan bijvoorbeeld betekenen dat verschillen in visie tussen eerstegraders en tweedegraders door de onderzoeksopzet onderbelicht zijn gebleven. Daar staat andermaal tegenover dat men in verwant onderzoek op dit punt tot dezelfde conclusies is gekomen (Hylkema, 1990; Hofman & Lugthart, 1991).

Literatuur

- Ax, J. (1985). *Planningsgedrag van leraren, een empirisch onderzoek naar de onderwijsplanning door leraren in het voortgezet onderwijs* (Academisch proefschrift). Lisse: Swets & Zeitlinger.
- Boyd, W. L., & Crowson, R. L. (1981). The changing conception and practice of public school administration. *Review of Research in Education*, 9, 311-373.
- Brandsma, H. P., & Stoel, W. G. R. (1986). *Schoolleiders in het voortgezet onderwijs. Een exploratieve studie naar determinanten en effecten van schoolleiderskenmerken*. Groningen: RION.
- Cohen, P. S., March, J. G., & Olson, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, 17, 1-25.
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 35, 15-24.
- Els, T. J. M., Extra, G., & Os, Ch. van (1977). *Handboek voor de Toegepaste Taalkunde. Het leren en onderwijzen van vreemde talen*. Groningen: Wolters-Noordhoff.
- Everitt, B. S. (1974). *Cluster analysis*. London: Heinemann.
- Firestone, W. A., & Wilson, B. L. (1985). Using bureaucratic and cultural linkages to improve instruction: The principals' contribution. *Educational Administration Quarterly*, 21(2), 7-30.
- Friebel, A. J. J. M. (1994). *Planning van onderwijs en het gebruik van planningsdocumenten: doet het ertoe?* (Academisch proefschrift). Oldenzaal: Dinkeldruk.
- Fullan, M. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Grift, W. van der (1987). *De rol van schoolleider bij onderwijsvernieuwingen* (Academisch proefschrift). 's-Gravenhage: VUGA-uitgeverij.
- Hanson, E. M. (1979). *Educational administration and organizational behavior*. Boston: Allyn & Bacon.
- Hofman, R. H., & Lugthart, E. (1991). *Interne capaciteit in het voortgezet onderwijs*. Groningen: RION.
- Hylkema, W. F. S. (1990). *Docenten en hun vaksecties* (Academisch proefschrift). Nijmegen: Katholieke Universiteit Nijmegen.
- Ingersoll, R. M. (1994). Organizational control in secondary schools. *Harvard Educational Review*, 64, 150-172.
- Kieser, A., & Kubicek, H. (1983). *Organisation*. Berlin/N. New York: Walter de Gruyter.
- Kremers, E. J. J. (Red.) (1990). *Overzicht van leerresultaten aan het eind van de eerste fase voortgezet onderwijs*. Arnhem: CITO.
- Krogt, F. J. van der (1983). *Probleemoplossing en beleidsvorming in scholen*. (Academisch proefschrift). Leiden: Rijksuniversiteit Leiden.
- Krogt, F. J. van der, & Oosting, J. (1990). Op zoek naar empirische configuraties in scholen. In J. H. G. I. Giesbers & F. J. van der Krogt (Red.), *Werken met modellen in organisaties* (pp. 89-100). Alphen a/d Rijn: Samsom H. D. Tjeenk Willink.
- Kuhlemeier, H., & Bergh, H. van den (1995). *Het belang van de vaksectie aardrijkskunde voor de aardrijkskundeprestaties*. Arnhem: Cito.

- Mahieu, P. (1988). *Participatie aan schoolmanagement. Een organisatiesociologische benadering*. (ongepubliceerde dissertatie). Antwerpen: UIA.
- Marwijk Kooy-von Baumhauer, L. van (1984). *Scholen verschillen, een verkennend vergelijkend onderzoek naar het intern functioneren van vijftwintig scholengemeenschappen VWO- HAVO- MAVO*. Groningen: Wolters-Noordhoff.
- Marx, E. C. H. (1975). *De organisatie van scholengemeenschappen in onderwijskundige optiek*. Groningen: Wolters-Noordhoff.
- Marx, E. C. H. (1984). Theoretisch overzicht van contingentiebenaderingen. In B. P. M. Creemers (Red.), *Handboek Schoolorganisatie en Onderwijsmanagement* (pp. 1410-1 – 1410-27). Alphen aan den Rijn: Samsom.
- Marx, E. C. H. (1987). Vermogen van scholen tot het voeren van bestuurlijk beleid. In L. Genemans (Red.), *Autonomie van scholen en deregulering* (pp. 17-26). Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Meijnen, G. W. (1987). Onderwijssociologie. In N. A. Lagerweij & J. F. Vos (Red.), *Onderwijskunde, een inleiding* (pp. 47-70). Groningen, Wolters-Noordhoff.
- Meyer, J. W., & Rowan, B. (1983). The structure of educational organizations. In J. Baldrige & T. Deal (Eds.), *The dynamics of organizational change in education*. California: Berkeley.
- Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Mintzberg, H. (1983). *Power in and around organizations*. Englewood Cliffs, NJ: NY: Prentice Hall.
- Murphy, J. (1988). Methodological, measurement, and conceptual problems in the study of instructional leadership. *Educational Evaluation and Policy Analysis*, 10 (2), 117-139.
- Pelkmans, A., & Vrieze, G. (1987). *Meer over coördinatie en management in het voortgezet onderwijs*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Rosenholtz, S. J. (1985). Effective schools: Interpreting the evidence. *American Journal of Education*, 93, 352-388.
- Rowan, B., Raudenbush, S. W., & Kang, S. J. (1989). *School climate in secondary schools: a multi level analysis*. Paper presented to the ESRC International Conference on Multilevel Methods in Educational Research, Edinburgh, August 1989.
- Scheerens, J., & Vilsteren, C. A. van (1988). De school als organisatie. *Onderwijskundig Lexicon, (II)*, pp. F 1100-1 – F 1100-29. Alphen a/d Rijn: Samsom.
- Scheerens, J. (1989). Onderzoek naar schoolorganisatorische effectiviteit en de contingentiebenadering. In J. Scheerens & J. C. Verhoeven (Red.), *Schoolorganisatie, beleid en onderwijskwaliteit* (pp. 9-21). Lisse: Swets & Zeitlinger.
- Schuit, J. P. J., & Slegers, P. (1990). De schoolorganisatiemodellen van Marx als alternatief voor de dichotome benadering van scholen. In J. H. G. I. Giesbers & F. J. van der Krogt (Red.), *Werken met modellen in organisaties* (pp. 75-86). Alphen a/d Rijn: Samsom H. D. Tjeenk Willink.
- Schuit, J. P. J. (1993). *Organisatieproblemen in het voortgezet onderwijs?: een studie naar de door docenten ervaren problemen met aspecten van de schoolorganisatie* (Academisch proefschrift). Nijmegen: Universiteitsdrukkerij.
- Siskin, L. S. (1991). Departments as different worlds: Subject subcultures in secondary schools. *Educational Administration Quarterly*, 17(2), 134-160.
- Siskin, L. S. (1994). *Realms of knowledge: academic departments in secondary schools*. London: Falmer Press.
- Slegers, P. J. C., Bergen, Th. C. M. & Giesbers, J. G. H. I. (1992). Schoolen beleidsvoering. *Pedagogische Studiën*, 69, 177-200.
- Smets, P. (1985a). De school als organisatie. In J. L. Peschar & A. A. Wesselingh (Red.), *Onderwijssociologie, een inleiding* (pp. 251-278). Groningen: Wolters-Noordhoff.
- Smets, P. (1985b). *Koördinatie en management in het voortgezet onderwijs*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Smets, P. (1986). Middenmanagement in het onderwijs. In A. M. L. van Wieringen (Red.), *Management van onderwijsinstellingen* (pp. 176-190). Groningen: Wolters-Noordhoff.
- Stoel, W. (1994). *De taakhoud, de taakomvang en de taakbelasting van schoolleiders in het basisonderwijs*. Enschede: OCTO, Universiteit Twente.
- Vilsteren, C. A. van (1984). De school als professioneelbureaucratische organisatie. In B. P. M. Creemers et al. (Red.), *Handboek schoolorganisatie en onderwijsmanagement* (pp. 1510-1 – 1510-27). Alphen a/d Rijn: Samsom.
- Vilsteren, C. A. van, & Visscher, A. J. (1987). Schoolwerkplanning: mogelijk in schoolorganisaties?. *Handboek voor Schoolorganisatie en -management* (pp. 6120-1 – 6120-24). Alphen a/d Rijn: Samsom.

- Vilsteren, C. A. van, & Witziers, B. (1990a). Vaksecties en het onderwijskundig leiderschap van schoolleiders. In J. Imants & W. Weijzen (Red.), *Organisatie van Onderwijsinstellingen* (pp. 61-70). Amsterdam/Lisse: Swets & Zeitlinger.
- Vilsteren, C. A., Witziers, B. (1990b). Vaksecties en schoolleiding; twee ontwikkelingsparameters van schoolorganisaties. In J. H. G. I. Giesbers & F. J. van der Krogt (Red.), *Werken met modellen in organisaties* (pp. 113-124). Alphen a/d Rijn: Samsom H.D. Tjeenk Willink.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Weick, K. E. (1984). Management of organizational change among loosely coupled elements. In P. Goodman (Ed.), *Change in organizations* (pp. 375-409). San Francisco: Jossey-Bass.
- Willower, D. I. (1979). Ideology and science in organization theory. *Educational Administration Quarterly*, 15(3), 20-42.
- Witziers, B., & Vilsteren, C. A. van (1990). Vaksecties en de coördinatie van hun vakonderwijs. In J. Imants & W. Weijzen (Red.), *Curriculum en Schoolorganisatie* (pp. 61-71). Amsterdam/Lisse: Swets & Zeitlinger.
- Witziers, B. (1992). *Coördinatie binnen scholen voor voortgezet onderwijs* (Academisch proefschrift). Enschede: Technische Universiteit Twente.
- Witziers, B. (1993). Coördinatie en de effectiviteit van het onderwijs op het vakgebied wiskunde. *Tijdschrift voor Onderwijsresearch*, 18, 279-295.
- Zwaan, A. H. van der (1984). *Ontwerp van organisatie-onderzoek, leerboek voor de praktijk: programmering, ontwerpvarianten en probleemstellingen*. Assen: Van Gorcum.

Manuscript aanvaard 30-10-1995.

Auteur

B. Witziers is als onderzoeker werkzaam bij het onderzoeksinstituut OCTO van de Universiteit Twente te Enschede. Tevens is hij als docent verbonden aan de vakgroep Onderwijsorganisatie en -management van de Faculteit Toegepaste Onderwijskunde in dezelfde plaats.

Adres: Universiteit Twente, vakgroep Onderwijsorganisatie en -management, Postbus 217, 7500 AE Enschede

Abstract

Coordination within secondary schools: departments and educational leadership

B. Witziers. *Pedagogische Studiën*, 1996, 73, 198-214.

Coordination of the curricular domain within secondary schools is the main theme of this article. Special attention is given to the coordinative role of departments and school management. The results, based on a survey including 376 teachers and 103 departments dispersed over 97 schools, show that departments, although there are differences between departments in different subject fields, are important units within secondary school with respect to the curricular domain. In this respect a portrait of departments emerges that is strikingly different from the dominant image of an egg-crate structure staffed by isolated teachers. Furthermore, the results indicate that the coordinative role of department covary with the extent of educational leadership within secondary schools.

De deugdenbenadering van de morele opvoeding

Een poging tot conceptuele verheldering

J. W. Steutel*

Samenvatting

In de afgelopen jaren is veel geschreven over de cultivering en ontwikkeling van deugden. Toch ontbreekt tot nu toe een systematische verheldering van de definiërende kenmerken van de zogenaamde deugdenbenadering van de morele opvoeding. In dit artikel worden drie definities van zo'n benadering onderscheiden en in een logische volgorde besproken. Bij elke definitie wordt onderzocht of deze al dan niet van toepassing is op de cognitieve ontwikkelingstheorie van Kohlberg. Betoogd wordt dat zelfs de vroege theorie van Kohlberg volgens de eerste twee definities als een deugdenbenadering moet worden beschouwd. Alleen op grond van een derde definitie waarin wordt gerefereerd aan een deugdethiek, kan de deugdenbenadering als een werkelijk andere en nieuwe conceptie van morele opvoeding worden opgevat.

Inleiding en doelstelling

Nog niet lang geleden, om precies te zijn in 1988, verscheen de tweede druk van *Can we teach children to be good?* In het tweede hoofdstuk van dat boekje geeft Straughan een heldere samenvatting van de voornaamste benaderingen van de morele opvoeding, waaronder waardenoverdracht, waardenverheldering en natuurlijk de welbekende cognitieve ontwikkelingstheorie. Wat echter opvalt is dat een belangrijke benadering ontbreekt of althans niet wordt genoemd, namelijk de zogeheten *deugdenbenadering*. Dat is merkwaardig, want

zelfs degene die de ontwikkelingen in het onderzoek naar de morele opvoeding op afstand volgt, zal het niet zijn ontgaan dat ongeveer vanaf het midden van de jaren tachtig een toenemend aantal publikaties is verschenen waarin termen als 'deugd', 'deugdzaam' en 'karakter' alleen al in de titel figureren. Volgens sommigen is er zelfs sprake van een nieuwe trend, van een verschuiving in de richting van een wezenlijk andere benadering van de morele opvoeding, ook al grijpt die benadering terug op inzichten van Aristoteles (vgl. b.v. Atherton, 1988; Nash, 1988; Nucci, 1989).

Maar ondanks het feit dat in de afgelopen periode veel gedacht en geschreven is over deugden en morele opvoeding, is er tot nog toe eigenlijk geen systematische poging ondernomen om te verhelderen wat onder een deugdenbenadering kan of moet worden verstaan. Toch is een uitgewerkt en onderbouwd antwoord op de vraag naar de kenmerken van een deugdenbenadering een voorwaarde om andere en wellicht meer wezenlijke kwesties doeltreffend te kunnen aanpakken. Te denken valt bij voorbeeld aan de vraag of de deugdenbenadering meer is dan oude wijn in nieuwe zakken en, als er inderdaad sprake is van een nieuwe, alternatieve benadering, hoe we deze dan moeten waarderen. Er is, kortom, behoefte aan verheldering van de specifieke aard van de deugdenbenadering en, in het verlengde daarvan, van de mogelijke verschillen tussen zo'n benadering en andere concepties van morele opvoeding.

In deze verhandeling wil ik een poging wagen aan deze behoefte tegemoet te komen. In de eerste plaats zal ik proberen drie definities van de term 'deugdenbenadering' te onderscheiden. Deze definities zijn geen willekeurige constructies mijnerzijds, maar expliciteringen van betekenissen die te vinden zijn in publikaties over deugden en morele opvoeding. Bij de bespreking van die definities zal ik een logische volgorde aanhouden: begonnen wordt met

* Graag wil ik Dwight Boyd, de leden van de TP-overleggroep aan de VU en de beoordelaars van *Pedagogische Studiën* bedanken voor hun constructieve commentaar op een eerdere versie van dit artikel.

de meest ruime definitie, waarna, in de definities die volgen, steeds condities zullen worden toegevoegd. Anders gezegd, de ordening loopt van de definitie met de minste inhoud en de meeste omvang tot de definitie met de meeste inhoud en de minste omvang, met als gevolg dat de deugdenbenadering zogezegd steeds 'radicaler' wordt.

In de tweede plaats zal ik bij elke definitie onderzoeken of deze al dan niet van toepassing is op de cognitieve ontwikkelingstheorie van Kohlberg en de zijnen. Zo'n vergelijkende analyse is niet alleen bedoeld om een toelichting te geven die ons inzicht in de voorgestelde definities kan verdiepen. Ook kan een dergelijke vergelijking duidelijk maken of de deugdenoptiek zich inderdaad onderscheidt van andere benaderingen en in dat opzicht wellicht een nieuwe ontwikkeling in het denken over morele opvoeding representeert. De theorieën van Kohlberg zijn lange tijd toonaangevend geweest in het onderzoek naar de morele opvoeding (en in zekere zin nog steeds dominant). Als de deugdenbenadering dus werkelijk iets anders en iets nieuws is, dan zal een plausibele definitie moeten worden geboden die in elk geval *niet* van toepassing is op de cognitieve ontwikkelings-theorie.

Een eerste definitie: karaktertrekken als doelen van morele opvoeding

In de opvoeding, inclusief de morele opvoeding, worden allerlei soorten doelen nagestreefd, van vaardigheden en capaciteiten tot attitudes en gezindheden, van vormen van kennis en inzicht tot gewoontes en gebruiken. In het voetspoor van Frankena (1973a, p. 20) zouden we al die opvoedingsdoelen *disposities* kunnen noemen. Maar we moeten dan wel beseffen dat we de term 'dispositie' in een brede, technische betekenis gebruiken, ongeveer als synoniem van het Engelse woord 'state of mind' of 'quality of mind'.

Als we de term op deze manier gebruiken, zijn deugden per definitie disposities die we op de een of andere manier positief waarderen. En ondeugden zijn dan vanzelfsprekend disposities die we op de een of andere wijze negatief

waarderen. Maar niet alle disposities die we positieve waarde toekennen zijn deugden. Neem bij voorbeeld taalvaardigheid, wiskundig inzicht, vitaliteit, geestigheid, levensvreugde en intelligentie. Stuk voor stuk eigenschappen die we normaal bewonderen of op een andere manier positief waarderen. Maar desondanks is de term 'deugd' niet van toepassing. Evenmin zijn alle disposities die we negatieve waarde toekennen ondeugden. Te denken valt aan onhandigheid, emotionele instabiliteit, een slecht geheugen, gebrek aan fantasie, verwardheid en muzikale ongevoeligheid. Over het algemeen zullen we deze eigenschappen betreuren of anderszins negatief beoordelen. Maar het zou vreemd zijn zulke disposities als ondeugden te kwalificeren.

Kennelijk vormen deugden en ondeugden een bepaalde subklasse van disposities. Een subklasse die in dit verband vaak wordt genoemd is de groep van *karaktertrekken*. En inderdaad, niet alleen zijn alle deugden en ondeugden typische karaktertrekken, ook lijkt het mogelijk alle karaktertrekken als deugden of ondeugden op te vatten. Wanneer we iemands karakter in positieve termen beschrijven, bij voorbeeld door te zeggen dat hij rechtvaardig, moedig, eerlijk, loyaal en behulpzaam is, dan geven we per definitie een opsomming van deugden. En wanneer we een negatieve beschrijving geven van iemands karakter, bij voorbeeld door hem partijdig, onbetrouwbaar, leugenachtig, laf of ongeduldig te noemen, dan refereren we per definitie aan ondeugden (vgl. Dent, 1984, pp. 9-10).

Welnu, volgens de eerste definitie die ik aan de orde wil stellen, is een benadering van de morele opvoeding een deugdenbenadering indien, en alleen indien, de cultivering van bepaalde karaktertrekken of deugden wordt nagestreefd. Aangezien opvoeden per definitie een doelgerichte activiteit is, worden in elke benadering van de morele opvoeding bepaalde disposities als opvoedingsdoelen aangemerkt. En wanneer tot die disposities ook karaktertrekken of deugden worden gerekend, is dat gegeven op zichzelf tegelijk noodzakelijk en voldoende om van een deugdenbenadering te kunnen spreken.

Deze eerste definitie is nogal ruim. Het is, denk ik, gemakkelijk aannemelijk te maken dat volgens deze definitie bepaalde concepties van

morele opvoeding die door Straughan (1988) worden onderscheiden als deugdenbenaderingen kunnen worden opgevat. Neem bij voorbeeld de benadering die bekend staat als waardenoverdracht ('value transmission'). Volgens deze aloude conceptie bestaat morele opvoeding in de overdracht van bepaalde waarden. Als voorbeelden noemt Straughan onder andere 'truthfulness', 'perseverance', 'patience' en 'respect for others' (o.c., pp. 13-14). En zou het doorgeven van zulke waarden iets anders kunnen zijn dan de cultivering van de corresponderende en gelijknamige deugden? Zelfs de benadering die expliciet het inprenten van waarden van de hand wijst, de zogenaamde waardenverheldering ('values clarification'), pleit impliciet voor de cultivering van bepaalde deugden. Zo wordt de leerling geleerd de opvattingen van anderen serieus te nemen en te respecteren, dat wil zeggen de deugd van verdraagzaamheid bijgebracht. Bovendien wordt de leerling aangemoedigd de kloof tussen de eigen waardenoriëntatie en het eigen gedrag te overbruggen, hetgeen impliceert dat belang wordt gehecht aan de ontwikkeling van zulke deugden als doorzettingsvermogen, zelfbeheersing en vasthoudendheid. Wellicht valt zelfs de stelling te verdedigen dat volgens de eerste definitie *alle* bekende benaderingen van de morele opvoeding deugdenbenaderingen zijn. Welke karaktertrekken als deugden of opvoedingsdoelen worden aangemerkt kan van benadering tot benadering variëren; maar *dat* de cultivering van bepaalde karaktertrekken wordt voorgestaan zou heel goed een gemeenschappelijk kenmerk kunnen zijn van alle benaderingen.

Geldt dit echter ook voor de cognitief-structuralistische benadering van Kohlberg? Hij is toch degene geweest die de deugdenbenadering denigrerend 'the bag of virtues approach' heeft genoemd? Om een antwoord op deze vragen te vinden is het van belang een onderscheid te maken tussen twee fasen in de ontwikkeling van Kohlbergs denken over de morele opvoeding. Kenmerkend voor de eerste fase is dat in feite slechts één opvoedingsdoel wordt erkend en verdedigd, namelijk het stimuleren van de stadiumgewijze ontwikkeling van de *structuur* van moreel redeneren. In de tweede fase, die ongeveer begint in het midden van de jaren zeventig, erkent Kohlberg dat ook

het bijbrengen van morele *inhouden* een belangrijke doelstelling is van morele opvoeding. In tegenstelling tot de eerste fase wordt de opvoeder niet langer uitsluitend gezien als een 'Socratic facilitator of structural development' maar bovendien als een 'socializer teaching value content and behavior' (Kohlberg, 1978, p. 84).

Als we de eerste, ruime definitie volgen, is de theorie van Kohlberg uit de tweede fase een duidelijk voorbeeld van een deugdenbenadering. Zoals bekend komt in die fase 'the just community approach' tot ontwikkeling. Daarin wordt gepoogd de theorie aangaande de ontwikkeling van structuren te integreren met een theorie aangaande de socialisatie van inhouden. En die morele inhouden worden als deugden uitgelegd. Positieve trekken die verband houden met 'a sense of community', zoals vertrouwen, zorg, verantwoordelijkheid, discipline en democratische participatie, worden als 'master virtues' bestempeld en functioneren als centrale opvoedingsdoelen (vgl. b.v. Power, Higgins & Kohlberg, 1989). Het is echter zeer de vraag of ook Kohlbergs theorie uit zijn eerste fase een deugdenbenadering kan worden genoemd. Met name in deze fase heeft hij scherpe kritiek geleverd op 'the character education approach' waarin de cultivering van 'a bag of virtues' wordt bepleit. Daartegenover plaatst hij zijn eigen theorie waarin alleen de ontwikkeling van structuren van moreel oordelen als opvoedingsdoel wordt erkend.

Toch valt op verschillende manieren aanneemelijk te maken dat zelfs in deze eerste fase Kohlbergs theorie in feite een deugdenbenadering is. Eén van die manieren is de volgende. Bij herhaling heeft Kohlberg de relatie tussen moreel oordelen en moreel gedrag aan de orde gesteld. Volgens hem blijkt uit onderzoek "that advance in moral judgment seems to correlate with more mature moral action" (Kohlberg, 1971, pp. 82-83). Zo is bijvoorbeeld aangetoond dat het percentage proefpersonen dat onder bepaalde omstandigheden bedrog pleegt lager wordt naarmate het stadium van morele ontwikkeling hoger is. Maar ook al is het morele oordeel wellicht de factor die de meeste invloed uitoefent op moreel gedrag, het is zeker niet de enige factor. Er is een kloof tussen oordelen en handelen, zelfs in het postconventionele stadium, en andere factoren zijn nodig om

die kloof te dichten. Tot die factoren rekent Kohlberg onder andere 'ego strength' (vgl. b.v. Kohlberg, 1971, pp. 79, 82-84; 1976, pp. 180-181). Deze technische, theoretische term heeft onder andere betrekking op datgene wat in de dagelijkse taal 'wilskracht' wordt genoemd. En wilskracht wordt niet alleen zelf als een deugd beschouwd, het is ook een factor die is opgebouwd uit deugden, waaronder zelfbeheersing, doorzettingsvermogen, geduld en matigheid (vgl. Steutel, 1992, hfdst. 2). Zo zien we dat Kohlberg ook in zijn eerste fase de ontwikkeling van deugden belangrijk vindt. Ook al gebruikt hij de term niet, de zogenaamde deugden van wilskracht worden noodzakelijk geacht om het morele oordeel in praktijk te brengen. En daarmee is zijn theorie, gegeven onze eerste definitie, een voorbeeld van een deugdenbenadering.

Een tweede definitie: alleen karaktertrekken als doelen van morele opvoeding

Men zou zich kunnen afvragen wat de waarde is van onze eerste definitie. Immers, als volgens deze definitie alle bekende benaderingen in feite deugdenbenaderingen zijn, inclusief die van Kohlberg in zijn eerste fase, dan mist de definitie elk differentiërend vermogen, met als gevolg dat evenmin op grond van deze definitie een nieuwe benadering in het denken over morele opvoeding kan worden aangewezen. Maar is dat een probleem? Ook al differentieert de definitie zelf niet, zij wijst ons wel een interessante weg om benaderingen te onderscheiden, namelijk aan de hand van verschillen in karaktertrekken die als opvoedingsdoelen worden aangemerkt. Hoe dan ook, als we geïnteresseerd zijn in een definitie die de deugdenbenadering onderscheidt van andere benaderingen, dan zullen we definiërende condities moeten toevoegen die werkelijk als *differentia specifica* fungeren. Welke condities zouden dat kunnen zijn?

Het feit dat bepaalde karaktertrekken als opvoedingsdoelen worden aangemerkt, is volgens de eerste definitie noodzakelijk en voldoende om te kunnen spreken van een deugdenbenadering. We zouden deze definitie

kunnen aanscherpen door van een deugdenbenadering te spreken indien, en alleen indien, *uitsluitend* karaktertrekken als doelen van de morele opvoeding worden beschouwd. Volgens de eerste definitie is het feit dat slechts één van de opvoedingsdoelen een karaktertrek is al voldoende om te kunnen spreken van een deugdenbenadering. Nu moeten echter *alle* disposities die als doelen worden aangemerkt karaktertrekken zijn. En dat is een aanzienlijke aanscherping. Want als we volgens deze tweede definitie een deugdenbenadering bepleiten, wijzen we niet alleen bepaalde karaktertrekken als opvoedingsdoelen aan, maar claimen we bovendien dat er buiten deze disposities niets de moeite waard is om in de morele opvoeding na te streven.

Een paar jaar geleden heb ik een verhandeling geschreven onder de titel *Deugden en morele opvoeding* (Steutel, 1992). In dat boekje verdedig ik een conceptie van morele opvoeding die, gegeven de tweede definitie, als een deugdenbenadering kan worden aangemerkt: alle doelen van de morele opvoeding worden begrepen in termen van deugden. Toen ik deze publikatie voorbereide ging ik ervan uit dat binnen de bekende benaderingen van de morele opvoeding, met uitzondering van 'the character education approach', doelen worden bepleit die als zodanig geen karaktertrekken zijn. Daarbij dacht ik met name aan de theorieën van Kohlberg. Maar klopt dat? Of is volgens onze tweede definitie zelfs de benadering van Kohlberg toch een deugdenbenadering?

Laten we, om te beginnen, Kohlbergs theorie uit zijn eerste fase bezien. Kenmerkend voor deze fase is dat slechts één opvoedingsdoel wordt erkend, namelijk de ontwikkeling van de vorm of de structuur van het moreel redeneren. Deze stadiumgewijze ontwikkeling moet uitmonden in een postconventionele moreel-oordeelsstructuur. Opmerkelijk is nu dat Kohlberg in verschillende publikaties uit zijn eerste fase heeft verdedigd dat iemand die overeenkomstig zo'n structuur oordeelt en handelt in feite drager is van de deugd van rechtvaardigheid (vgl. m.n. Kohlberg, 1970; 1971, pp. 80-81). Het doel van de morele opvoeding bestaat volgens hem dus niet in een 'zak met deugden', zoals binnen 'the character education approach' het geval is, maar ten leste slechts in één deugd, de complexe en abstracte

deugd van rechtvaardigheid. En dat zou, gegeven onze tweede definitie, zijn theorie tot een deugdenbenadering maken, aangezien naast deze deugd geen andere disposities als opvoedingsdoelen worden erkend.

Men zou kunnen tegenwerpen dat binnen een echte deugdenbenadering altijd *meerdere* karaktertrekken als opvoedingsdoelen worden aangemerkt en dat het dan ook niet toevallig is dat onze tweede definitie in meervoud is gesteld. Maar zelfs wanneer we deze tegenwerping plausibel vinden, dan nog blijft Kohlbergs theorie uit zijn eerste fase een deugdenbenadering. Want de deugd van rechtvaardigheid, zoals door Kohlberg uitgelegd, impliceert tal van andere deugden, waarbij onder andere te denken valt aan zogenaamde *intellectuele* deugden, zoals openheid van geest, tolerantie tegenover rivaliserende opvattingen en respect voor rationele argumenten. Iemand die rechtvaardig is, aldus Kohlberg, is in staat te oordelen overeenkomstig morele principes, in het bijzonder de principes van rechtvaardigheid. Dergelijk oordelen heeft niet alleen individuele gedragingen tot onderwerp, maar ook regels en wetten (vgl. b.v. Kohlberg, 1971, p. 59). Het zal duidelijk zijn dat zo'n toetsing van gedragingen, regels en wetten aan principes van rechtvaardigheid over het algemeen een veeleisende en ingewikkelde aangelegenheid is. Om dergelijke complexe oordeelsprocessen tot een goed einde te brengen, hebben we de intellectuele deugden nodig. Want de praktisering van deze deugden biedt een zekere garantie dat onze oordeelsvorming (inclusief ons oordelen over de rechtvaardigheid van gedragingen, regels en wetten) resulteert in opvattingen die juist zijn of althans op goede gronden berusten (vgl. Steutel & Spiecker, 1995).

Kohlberg verwacht van de drager van de deugd van rechtvaardigheid de competentie om adequaat te kunnen oordelen over kwesties van rechtvaardigheid. Als we zo'n oordeelscompetentie van waarde vinden, zo hebben we betoogd, dan zullen we tevens waarde moeten hechten aan een keur van intellectuele deugden. De drager van de deugd van rechtvaardigheid wordt echter ook gekenmerkt door een *handelingscompetentie*. Naast adequaat oordelen is zo'n persoon in staat zijn doen en laten op dat oordeel af te stemmen. En we hebben in de voorgaande paragraaf gezien dat voor het dich-

ten van de kloof tussen oordelen en handelen de deugden van wilskracht nodig zijn. Met andere woorden, die ene deugd van rechtvaardigheid, zoals uitgelegd door Kohlberg, blijkt zowel intellectuele deugden als deugden van wilskracht te omvatten. Als dat geen flinke zak met deugden is²! In Kohlbergs tweede fase wordt die zak nog verder gevuld. Dan rekent hij immers ook het aanleren van morele inhouden tot de morele opvoeding. En die inhouden worden binnen 'the just community approach' uitgelegd als deugden, in hoofdzaak als karaktertrekken die constitutief zijn voor 'a love of the common life'.

Op grond van deze analyse kunnen we waarschijnlijk de conclusie trekken dat het differentiërend vermogen van de tweede definitie niet veel groter is dan dat van de eerste definitie. Indien ook volgens onze tweede definitie zelfs de vroege theorie van Kohlberg een deugdenbenadering is, mag in elk geval worden aangenomen dat zo'n benadering bepaald geen alternatieve en nieuwe ontwikkeling in het denken over morele opvoeding representeert. Misschien ligt deze conclusie ook wel voor de hand. In het onderzoek naar de morele opvoeding staat de vraag naar de disposities van de morele persoon per definitie in het centrum van de belangstelling. Deze disposities vormen immers de doelen van de morele opvoeding. En is het niet vanzelfsprekend om zulke disposities op te vatten als deugden, of ten minste als componenten of aspecten van deugden?

Een derde definitie: morele opvoeding die gefundeerd is in een deugdedethiek

Toch kan nog steeds niet worden uitgesloten dat de deugdenbenadering zich op wezenlijke punten onderscheidt van de cognitieve ontwikkelingstheorie. Zeker, als we het denken over deugden als zodanig als criterium hanteren, kunnen ook de opvattingen van Kohlberg als een deugdenbenadering worden aangemerkt. Maar het zou kunnen zijn dat in recente publikaties op het vlak van de morele opvoeding *op een andere manier* over deugden wordt nagedacht. En misschien moeten we een derde definitie introduceren die louter betrekking heeft

op die recente, alternatieve benadering *binnen* het denken over deugden.

Vanzelfsprekend zal zo'n definitie de touwtjes nog wat strakker moeten aantrekken. Immers, op grond van deze definitie zullen we die alternatieve benadering moeten kunnen onderscheiden van de cognitieve ontwikkelings-theorie. En aangezien die differentiatie met de eerste twee definities niet lukte, zullen we een aangescherpte definitie moeten formuleren door één of meer condities toe te voegen. Hoe zou deze definitie er uit kunnen zien?

Een interessante kandidaat is naar mijn idee een definitie waarin een verband wordt gesticht tussen een deugdenbenadering en een deugd-ethiek. Precieser gezegd: van een deugdenbenadering is volgens deze definitie sprake indien, en alleen indien, de karaktertrekken die als opvoedingsdoelen zijn aangewezen worden gefundeerd in een ethiek van deugd. Maar wat is precies een deugdeethiek? En wat betekent het om de doelen van de morele opvoeding in zo'n ethiek te funderen?

Vaak wordt gewezen op de recente herleving van het denken over deugden en ondeugden in de wijsgerige ethiek, en dan vooral in de Angelsaksische wereld. En niet ten onrechte: in de afgelopen decennia zijn binnen deze tak van de wijsbegeerte talrijke artikelen geschreven over de aard van deugden in het algemeen, over verschillende typen van deugden, over deugden en karakter, deugden en emoties, deugden en motieven, alsmede over de vraag naar de mogelijke eenheid van de deugden en de relatie tussen deugden en het menselijk floreren. Zulke systematische reflectie op deugden is echter nog niet identiek aan de ontwikkeling van een deugdeethiek. In tegenstelling tot veel ander wijsgerig onderzoek naar deugden, wordt zo'n ethiek gepresenteerd als een *alternatief* voor of een *concurrent* van de standaardconcepties in de wijsgerige ethiek, te weten het kantianisme, het utilisme en het intuïtionisme. Neem bijvoorbeeld het recente boek van Rawls, met als titel *Political Liberalism* (1993). In dat werk bespreekt Rawls de aard en functie van politieke deugden, zoals tolerantie, wederzijds respect, redelijkheid, burgerzin en compromisbereidheid. Als zodanig draagt hij bij aan de herleving van de filosofische reflectie op deugden. Maar omdat zijn analyse van de politieke deugden deel uitmaakt van een ethi-

sche theorie die moet worden opgevat als een variant van het kantianisme, is hij duidelijk geen verdediger van een deugdeethiek. En de vraag is: waarin onderscheidt een ethiek van deugd zich van deze welbekende andere wijsgerige concepties?

In de relevante literatuur worden meestal twee specifieke kenmerken van een deugdethiek genoemd (vgl. m.n. Frankena, 1970, pp. 6-11; 1973b, pp. 63-65; 1973c; Brandt, 1981, pp. 271-273; Loudon, 1984; Watson, 1990; Slote, 1992, pp. 88-93). Een eerste criterium is dat binnen zo'n ethiek aretaïsche concepten en oordelen een basale of fundamentele plaats krijgen toegewezen, althans in verhouding tot deontische concepten en oordelen. Voorbeelden van aretaïsche termen (van *arete*, Grieks voor 'uitmuntende eigenschap') zijn 'goed' en 'slecht', 'moedig' en 'laf', 'deugd' en 'ondeugd'. Zulke termen kunnen worden onderscheiden van zogenoemde deontische termen (van het Griekse *deon*, dat vaak met 'plicht' wordt vertaald), waarvan 'juist' en 'verkeerd', 'plicht' en 'verplichting', 'behoren' en 'moe-ten' bekende voorbeelden zijn. De laatste groep van termen is uitsluitend van toepassing op daden en praktijken, terwijl de eerstgenoemde groep ook betrekking heeft op intenties, motieven, karaktertrekken en personen. Bovendien: in tegenstelling tot deontische termen, refereren aretaïsche termen aan scalaire kwaliteiten, dat wil zeggen het zijn adjectieven met comparatieven (vergroten trap) en superlatieven (overtreffende trap). Als we bij voorbeeld zeggen dat iets goed of slecht is, verwijzen we naar scalaire eigenschappen, aangezien we kunnen spreken van 'het betere' en 'het beste', of 'het slechtere' en 'het slechtste' (vgl. Urmson, 1968, p. 92).

Wat nu typerend is voor een deugdeethiek is dat aan aretaïsche concepten en oordelen een primaire of basale plaats wordt toegekend, terwijl deontische begrippen en oordelen of als onnodig dan wel inadequaat worden beschouwd, of gezien worden als afgeleid uit of te reduceren tot aretaïsche concepten en oordelen. Een voorbeeld van de eerste variant (deontische concepten zijn onnodig of inadequaat) is verdedigd door Anscombe (1958). In haar baanbrekende en tegelijk controversiële artikel 'Modern moral philosophy' verwerpt ze alle begrippen van morele plicht en verplichting

omdat volgens haar zulke begrippen buiten het kader van een goddelijke wet niet intelligibel zijn. Haars inziens kunnen we beter een Aristotelische ethiek ontwikkelen waarin de ethische concepten en oordelen uitsluitend aretaïsch zijn. Slote heeft echter een voorbeeld ontwikkeld van de tweede variant van een deugdetiek (deontische concepten zijn afgeleid of reduceerbaar). Hij kent een basale plaats toe aan het aretaïsche begrippenpaar 'bewonderenswaardig'/'betreurenswaardig', dat hij zowel van toepassing verklaart op daden als motieven en karaktertrekken (Slote, 1992, pp. 95-96). En volgens hem kunnen imperatieven en 'behoort'-oordelen worden afgeleid uit aretaïsche oordelen waarin van dat begrippenpaar gebruik wordt gemaakt (vgl. pp. 159-167).

Een tweede criterium dat vaak wordt genoemd is dat in een ethiek van deugd aretaïsche oordelen aangaande daden worden herleid tot aretaïsche oordelen over personen en hun karakter. Een moreel goede daad wordt dan bij voorbeeld gedefinieerd als gedrag dat een toonbeeld is van morele deugd, of een daad wordt als edelmoedig bestempeld als het een daad is die de edelmoedige persoon zou voltrekken. De grondgedachte is dat wanneer we daden in aretaïsche termen beoordelen, we zulks altijd doen tegen de achtergrond van de motieven, disposities of karaktertrekken die in deze daden tot uitdrukking komen. Met andere woorden, volgens het eerste criterium is een deugdetiek een *aretaïsche* ethiek (en geen deontische ethiek); en volgens het tweede criterium is een deugdetiek een specifiek soort aretaïsche ethiek, namelijk een aretaïsche *actorethiek* (en geen aretaïsche daadethiek).

Volgens onze derde definitie is een deugdenbenadering gefundeerd in een ethiek van deugd. Dat betekent in feite dat alle doelen van de morele opvoeding kunnen worden beschreven met behulp van aretaïsche oordelen, en dat deze oordelen of deel uitmaken van de kern van een ethiek van deugd, of uit zo'n funderende kern kunnen worden afgeleid. In een deugdenbenadering worden uitsluitend karaktertrekken als doelen van de morele opvoeding aangegeven. Als zodanig worden die karaktertrekken getypeerd in aretaïsche termen, namelijk als goede of bewonderenswaardige eigenschappen, of kortweg als deugden. En wanneer deze aretaïsche oordelen in ethisch opzicht als ba-

saal worden beschouwd, of als afgeleid uit zulke basale oordelen, dan wordt daarmee de morele opvoeding gefundeerd in een ethiek van deugd en is er sprake van een deugdenbenadering.

Volgens deze definitie is de theorie die ik in mijn *Deugden en morele opvoeding* heb ontwikkeld nadrukkelijk geen deugdenbenadering. In dat boekje heb ik verdedigd dat morele deugden zijn opgebouwd uit gezindheden om in overeenstemming met morele regels of principes te handelen. Sommige moreel-goede trekken, die ik non-teleologische deugden heb genoemd, corresponderen louter met morele regels die vastleggen welk gedrag moreel juist of moreel vereist is (zgn. plichten en verplichtingen). Deze opvatting is in strijd met het eerste criterium van een deugdetiek: aretaïsche oordelen (b.v. 'Trouw is een non-teleologische deugd') worden afgeleid uit deontische oordelen (b.v. 'Men behoort zijn beloften na te komen'). Omdat het onze plicht is bepaalde dingen te doen of te laten, is de gezindheid om dat gedrag ten toon te spreiden een deugd. Andere moreel-goede karaktereigenschappen, door mij teleologische deugden genoemd, corresponderen altijd ook en soms zelfs uitsluitend met morele regels die vastleggen welk gedrag moreel wenselijk is (zgn. *supererogativa*). Hier verdedig ik een opvatting die niet voldoet aan het tweede criterium van een deugdetiek: aretaïsche oordelen aangaande karaktertrekken (b.v. 'Behulpzaamheid is een teleologische deugd') worden afhankelijk gemaakt van aretaïsche oordelen aangaande daden (b.v. 'Het bieden van deze vorm van hulp is moreel wenselijk'). Omdat bepaald gedrag moreel wenselijk is, is de gezindheid om dat gedrag ten uitvoer te brengen een teleologische deugd. Kortom, in mijn boek fundeer ik de doelen van de morele opvoeding niet in een deugdetiek maar in een principe-ethiek. En aangezien ik dacht die principes te kunnen rechtvaardigen in termen van het grootste geluk voor het grootste aantal, heb ik een principe-ethiek van utilistische snit verdedigd (vgl. Steutel, 1992, pp. 122-123).

Maar wat te denken van de theorieën van Kohlberg? Is zijn conceptie van morele opvoeding volgens de derde definitie wel een deugdenbenadering? Deze vraag moeten we, naar ik meen, ontkennend beantwoorden. Van meet af

aan heeft Kohlberg zijn theorie gesitueerd binnen de deontologische of kantiaanse traditie (vgl. b.v. Kohlberg, 1976, pp. 182-183). En binnen deze traditie ziet men de funderende kern van de ethiek als opgebouwd uit deontische begrippen en oordelen of, in de woorden van Rawls, uit 'the concept of the right' en 'principles of the right'.

We hebben boven gezien dat Kohlberg een (of de) centrale doelstelling van de morele opvoeding typeert in termen van de deugd van rechtvaardigheid. Aretaïsche oordelen over de morele persoon en zijn karakter zijn dan ook geenszins afwezig in zijn theorie. Maar deze oordelen zijn nadrukkelijk niet primair of basaal. De aretaïsche evaluatie van de persoon en zijn karakter wordt afhankelijk gemaakt van deontische oordelen, in het bijzonder van principes van rechtvaardigheid. Om te bepalen welk gedrag moreel juist is, moeten we volgens Kohlberg een beroep doen op zulke principes. En de deugd van rechtvaardigheid is niets meer of minder dan het vermogen en de bereidheid om in oordelen en handelen deze principes in acht te nemen. Kohlberg, met andere woorden, typeert de deugd van rechtvaardigheid in termen van principes van rechtvaardigheid. En daarmee fundeert hij de morele opvoeding niet in een deugdedethiek maar in een principe-ethiek, en wel een principe-ethiek van deontologische signatuur.

Het laat zich aanzien dat onze derde definitie inderdaad differentieert tussen een deugdenbenadering en bekende andere benaderingen van de morele opvoeding. In elk geval worden bepaalde benaderingen buitengesloten, waaronder de theorie van Kohlberg. Maar mogen we ook aannemen dat de definitie recent ontwikkelde concepties van morele opvoeding insluit? De klassieke ethische theorie van Aristoteles, alsmede die van de epicuristen en de stoïcijnen, worden op goede gronden opgevat als voorbeelden van een deugdedethiek. In het verlengde daarvan kunnen we de ideeën over opvoeding die in deze theorieën worden verdedigd beschouwen als voorbeelden van een deugdenbenadering. Maar zijn er eigenlijk wel hedendaagse opvattingen over morele opvoeding die volgens de derde definitie deugdenbenaderingen zijn?

Het is niet eenvoudig op deze vraag een eenduidig antwoord te geven, met name niet omdat

in de meeste recente publikaties over deugden en morele opvoeding meta-ethische reflecties over het verschil tussen een plichtethiek en een deugdedethiek volledig ontbreken. Toch kunnen we, denk ik, de gestelde vraag bevestigend beantwoorden. Om te beginnen kan wellicht de morele opvoedingstheorie van White (1990) als een deugdenbenadering worden aangemerkt. Dat is althans mogelijk in het licht van zijn kritiek op 'law-based conceptions of morality' en zijn verdediging van een alternatief dat op de klassieken is geënt. Een duidelijker voorbeeld van een deugdenbenadering, in de derde betekenis van het woord, vinden we in een publikatie van Crittenden (1990). Ofschoon hij niet expliciet verwijst naar de criteria van een deugdedethiek, verdedigt hij naar eigen zeggen een Aristotelische benadering van moraliteit en morele ontwikkeling. Het meest duidelijke en uitgewerkte voorbeeld van een deugdenbenadering vinden we echter in een boek van Carr (1991). In deze verhandeling contrasteert hij een kantiaans perspectief op moraliteit ('an ethics of obligation') met een Aristotelische visie op het morele leven ('an ethics of virtue') (p. 87). En hij plaatst zijn eigen theorie van morele opvoeding nadrukkelijk in de laatstgenoemde traditie.

Ik geef ogenblikkelijk toe dat deze typering van recent gepubliceerde ideeën nadere uitleg en ondersteuning verdient. Maar als ik gelijk heb, valt er toch een benadering aan te wijzen die duidelijk anders is dan de cognitieve ontwikkelingstheorie. Dan is er inderdaad sprake van een nieuwe ontwikkeling die gebaseerd is op eeuwenoud gedachtengoed.

Noten

1. In de eerste fase van zijn denken over morele opvoeding formuleert Kohlberg verschillende punten van kritiek op de 'bag of virtues approach' (vgl. b.v. Kohlberg, 1970, pp. 31-37). Opvallend is dat hij in de tweede fase van zijn denken deze kritiek niet opnieuw aan de orde stelt, terwijl sommige punten toch ook betrekking hebben op zijn 'just community approach'. Denk bijvoorbeeld aan het punt dat volgens Kohlberg in empirisch onderzoek geen stabiele en situatie-overstijgende trekken gevonden worden, of aan het punt dat volgens hem elke lijst van deugden arbi-

trair of ethisch relatief is.

2. De opvatting dat Kohlbergs doel van morele opvoeding verbonden is met een hele groep van deugden is bepaald niet origineel. Zo heeft Boyd (1989) laten zien dat stadium zes in Kohlbergs theorie correspondeert met zulke deugden als 'benevolence', 'sympathy', 'humility' en 'a sense of responsibility'. Al veel eerder kwam Peters tot de conclusie dat het postconventionele niveau van moreel denken correleert met "quite a formidable 'bag of virtues'" (1981, p. 178). Hun analyse en argumentatie is echter anders dan de mijne.

Literatuur

- Anscombe, G.E.M. (1958). Modern moral philosophy. *Philosophy*, 33, 1-19.
- Atherton, J.M. (1988). Virtues in moral education: Objections and replies. *Educational Theory*, 38, 299-310.
- Boyd, D. (1989). The character of moral development. In L.P. Nucci (Ed.), *Moral development and character education. A dialogue* (pp. 95-123). Berkeley: McCutchan Publishing Corporation.
- Brandt, R.B. (1981). W.K. Frankena and ethics of virtue. *The Monist*, 64, 271-292.
- Carr, D. (1991). *Educating the virtues. An essay on the philosophical psychology of moral development and education*. London: Routledge.
- Crittenden, P. (1990). *Learning to be moral. Philosophical thoughts about moral development*. New Jersey: Humanities Press International.
- Dent, N.J.H. (1984). *The moral psychology of the virtues*. Cambridge: Cambridge University Press.
- Frankena, W.K. (1970). Prichard and the ethics of virtue. Notes on a footnote. *The Monist*, 54, 1-17.
- Frankena, W.K. (1973a). The concept of education today. In J.F. Doyle (Ed.), *Educational judgments. Papers in the philosophy of education* (pp. 19-32). London: Routledge & Kegan Paul.
- Frankena, W.K. (1973b). *Ethics*. Englewood Cliffs, NJ: Prentice-Hall.
- Frankena, W.K. (1973c). The ethics of love conceived as an ethics of virtue. *The Journal of Religious Ethics*, 1, 21-36.
- Kohlberg, L. (1970). Education for justice: a modern statement of the socratic view. In L. Kohlberg, *Essays on moral development. Vol. I: The philosophy of moral development. Moral stages and the idea of justice* (pp. 29-48). San Francisco: Harper & Row, 1981.
- Kohlberg, L. (1971). Stages of moral development as a basis for moral education. In B. Munsey (Ed.), *Moral development, moral education, and Kohlberg. Basic issues in philosophy, psychology, religion, and education* (pp. 15-98). Birmingham, AL: Religious Education Press.
- Kohlberg, L. (1976). The cognitive-developmental approach to moral education. In D. Purpel & K. Ryan (Eds.), *Moral education. ...It comes with the territory* (pp. 176-195). Berkeley: Phi Delta Kappa.
- Kohlberg, L. (1978). Revisions in the theory and practice of moral development. In W. Damon (Ed.), *Moral development* (pp. 83-87). San Francisco: Jossey-Bass Inc.
- Louden, R.B. (1984). On some vices of virtue ethics. *American Philosophical Quarterly*, 21, 227-236.
- Nash, R.J. (1988). The revival of virtue in educational thinking: a postliberal appraisal. *Educational Theory*, 38, 27-39.
- Nucci, L.P. (Ed.) (1989). *Moral development and character education. A dialogue*. Berkeley: McCutchan Publishing Corporation.
- Peters, R.S. (1981). *Moral development and moral education*. London: George Allen & Unwin.
- Power, C., Higgins, A., & Kohlberg, L. (1989). The habit of the common life: building character through democratic community schools. In L.P. Nucci (Ed.), *Moral development and character education* (pp. 125-143). Berkeley: McCutchan Publishing Corporation.
- Rawls, J. (1993). *Political liberalism*. New York: Columbia University Press.
- Slote, M. (1992). *From morality to virtue*. New York: Oxford University Press.
- Steutel, J.W. (1992). *Deugden en morele opvoeding. Een wijsgerig-pedagogische studie*. Meppel/Amsterdam: Boom.
- Steutel, J.W., & Spiecker, B. (1995). Rationele passies en intellectuele deugden. *Algemeen Nederlands Tijdschrift voor Wijsbegeerte*, 87, 1-14.
- Straughan, R. (1988). *Can we teach children to be good? Basic issues in moral, personal and social education*. Milton Keynes: Open University Press.
- Urmson, J.O. (1968). *The emotive theory of ethics*. London: Hutchinson University Library.
- Watson, G. (1990). On the primacy of character. In O. Flanagan & A.O. Rorty (Eds.), *Identity, character, and morality. Essays in moral psychology* (pp. 449-469). Cambridge, MA: The MIT Press.
- White, J. (1990). *Education and the good life. Beyond the National Curriculum*. London: Kogan Page.

Auteur

J.W. Steutel is als hoofddocent verbonden aan de vakgroep Pedagogiek van de Vrije Universiteit.

Adres: Van der Boechorststraat 1, 1081 BT Amsterdam.

Abstract

The virtue approach to moral education. Some conceptual clarifications

J.W. Steutel. *Pedagogische Studiën*, 1996, 73, 215-224.

There is a lot of talking and writing on education and virtues nowadays. In spite of this, a clear and convincing account of the defining characteristics of the virtue approach to moral education is still lacking. This paper suggests and discusses three different definitions of such an approach. With reference to each definition it is examined whether the virtue approach can be distinguished from other main approaches to moral education, in particular from the so-called cognitive-developmental approach. It is argued that only a definition that refers to an ethics of virtue will have the required differentiating capacity.

Internationale publikaties en kwaliteit

Een onderzoek naar het publikatiegedrag van Nederlandse pedagogen en onderwijskundigen*

P. M. Kroonenberg en R. van der Veer

Samenvatting

In dit artikel wordt betoogd dat het aantal internationale tijdschriftpublicaties geproduceerd binnen een project een goede aanduiding vormt voor de kwaliteit ervan. Op basis van door de universiteiten aangeleverde gegevens wordt een overzicht gepresenteerd van de (inter)nationale publikaties van Nederlandse pedagogen en onderwijskundigen. Aan de hand van dit overzicht wordt een aantal bevindingen van de 'VSNU-visitatiecommissie-Knoers' van kritische kanttekeningen voorzien.

Ruim tien jaar geleden vermeldde Van IJzendoorn en Van der Veer (1985) de resultaten van een beknopt onderzoek naar de internationale publikaties van Nederlandse pedagogen en onderwijskundigen. Gebruikmakend van de Corporate Index van het Institute of Scientific Information (ISI), een bestand waarin vermeld wordt hoeveel publikaties een instelling produceert, stelden zij vast dat Nederlandse onderzoekers verbluffend weinig in buitenlandse tijdschriften publiceerden. Van IJzendoorn en Van der Veer riepen hun collega's op hun leven te beteren temeer daar "steeds vaker gesproken wordt over zogenaamde visitatiecommissies

grotendeels samengesteld uit buitenlandse experts, die Nederlands onderzoek en onderwijs zouden moeten doorlichten" (Van IJzendoorn & Van der Veer, 1985, p. 6). De auteurs betoogden dat dergelijke commissies wel eens een groot gewicht zouden kunnen toekennen aan internationale publikaties als indicator voor de kwaliteit van het verrichte onderzoek.

Intussen is deze voorspelling bewaarheid geworden en het lijkt interessant na te gaan of en in hoeverre het publikatiegedrag van de Nederlandse pedagogen (en onderwijskundigen) veranderd is. Internationale publikaties zullen uitgangspunt en toetssteen zijn voor de meeste van de oordelen in dit artikel en daarom zullen we eerst beknopt aangeven waarom het niet onredelijk is het aantal internationale publikaties van een onderzoeker als een aanduiding voor zijn onderzoekskwaliteiten te zien. Nadat we vervolgens onze kwantitatieve resultaten gepresenteerd hebben, zullen we ten slotte nagaan in hoeverre onze resultaten sporen met de bevindingen van de 'commissie Knoers' die onlangs in opdracht van de VSNU een rapport over de kwaliteit van het Nederlands pedagogisch en onderwijskundig onderzoek het licht deed zien (*Quality Assessment of Research. Pedagogical and Educational Sciences*, 'Commissie Knoers', 1995).

Internationale publikaties als kwaliteitsaanduiding

Niemand zal willen beweren dat internationale publikaties het enige kwaliteitskenmerk zijn of dat hoge aantallen internationale publikaties kwaliteit garanderen, maar in het algemeen zal er een relatie bestaan tussen het aantal internationale publikaties van een onderzoeker of onderzoeksgroep en de kwaliteit van het verrichte onderzoek. Waarom is het wetenschappelijk gezien nastrevenswaardig internationaal te publiceren en waarom dienen wij internationale publikaties in het algemeen wat hoger aan te slaan dan nationale?

Ten eerste (zie ook Van IJzendoorn & Van der Veer, 1985) wordt het forum van discussie bij een internationale publikatie enorm uitgebreid. Voor de voortgang van de wetenschap is het essentieel dat men zijn gedachten blootstelt aan kritiek en deze kritiek kan des te beter uit-

* Exemplaren van het interne rapport dat aan dit artikel ten grondslag lag en het databestand zijn opvraagbaar bij de eerste auteur (email: kroonenb@rulfsw.leidenuniv.nl).

geoeft worden naarmate het aantal deskundigen dat in principe van een gedachte kennis kan nemen groter is. De redactie van een vooraanstaand internationaal tijdschrift beschikt, in principe, over een groter aantal deskundige beoordelaars dan een nationaal tijdschrift. Internationale publikaties (boeken, tijdschriften) worden in het algemeen door meer deskundigen gelezen. Door internationaal te publiceren maakt men zich dus kwetsbaarder en dit is gunstig voor de voortgang van de wetenschap.

Ten tweede (Van IJzendoorn & Van der Veer, 1985) geldt ook in de wetenschap de eis van wederkerigheid. Nederlands onderzoek leunt vaak zwaar op buitenlands onderzoek en indien men dit buitenlandse onderzoek kritisch verwerkt, dient men de buitenlandse onderzoekers in principe van deze kritiek op de hoogte te stellen. Indien dit niet gebeurt, kan er geen open kritische dialoog plaatsvinden en wordt het internationale wetenschappelijke forum niet in staat gesteld zijn werk te doen. Men kan slechts van kritiek profiteren als men er van in kennis wordt gesteld. Bovendien lijkt het ook niet eerlijk het buitenland onze diepe inzichten te onthouden. In de wetenschap geldt het zogeheten communaliteitsbeginsel (Merton, 1973), hetgeen we in dit geval zouden kunnen weergeven met het gebod: "Gij zult Uw collega's in kennis stellen van Uw jongste bevindingen". Ten slotte geldt dat de voortgang in de wetenschap gebaat is bij pluriformiteit. Op de wetenschappelijke markt mogen onze exotische inzichten niet ontbreken en juist als men van mening is dat het in internationale wetenschappelijke tijdschriften gepresenteerde onderzoek eenzijdig (positivistisch, empiristisch, a-theoretisch, etc.) van aard is, dan dient men in het belang van de wetenschap de eigen afwijkende inzichten met kracht aan de man te brengen. Het tegenargument dat voor bepaald wetenschappelijk onderzoek eenvoudigweg geen internationale belangstelling bestaat, lijkt ons niet sterk. Typerend voor *wetenschappelijk* onderzoek is nu juist dat het een boven-lokale strekking heeft die aan een van de redacties van de nu al meer dan 100 opvoedkundige of onderwijskundige ISI-tijdschriften duidelijk te maken moet zijn.

Bovenstaande argumenten impliceren dat een internationale gedachtenwisseling in principe gewenst en vruchtbaar is. Nu zou men nog

kunnen beweren dat het loutere feit dat iemand er in slaagt aan dit internationale debat deel te nemen nog niets zegt over de kwaliteit van zijn bijdrage, met andere woorden, dat internationale produktiviteit in het geheel niet samenhangt met kwaliteit. Dit argument lijkt ons niet geheel overtuigend. Immers, vooraanstaande internationale tijdschriften beschikken in principe over de meest deskundige beoordelaars (omdat zij uit een groter reservoir van beoordelaars kunnen kiezen). Ook hanteren zij een hoger afwijzingspercentage dan lokale tijdschriften. Dit betekent dat iemand die er in slaagt internationaal te publiceren in het algemeen een zware intercollegiale toets heeft doorstaan. Iets soortgelijks geldt voor publikaties bij gerenommeerde internationale uitgevers die veelal over zeer deskundige lezers beschikken die manuscripten of boekvoorstellen op hun waarde kunnen schatten. Internationale publikaties zullen dus door de bank genomen van voldoende kwaliteit zijn en van een hogere kwaliteit zijn dan nationale. Derhalve is het niet onredelijk bij de beoordeling van de kwaliteit van een onderzoeksproject de produktiviteitsmaat van het aantal internationale publikaties zwaar te laten wegen.

Op gevaar af de wetenschap te profaniseren en de metafoor van de markt verwisselend voor die van de sport, kan men dit ook als volgt inzien. Velen zijn in staat schaak- of fierljeppkampioen van het dorp te worden en sommigen zullen zich op grond van dit resultaat onverslaanbaar wanen. Maar het is duidelijk dat hun ware kwaliteiten pas blijken bij de interregionale kampioenschappen en dat niet iedereen aan de IOC-normen kan voldoen.

Overigens zij hier, wellicht ten overvloede, vermeld dat wij niet willen stellen dat (inter)nationale produktiviteit de enige maat is voor iemands kwaliteiten als onderzoeker. Uiteraard vallen er nog veel andere kwaliteitscriteria voor onderzoekers te bedenken (vgl. Beekman, 1986). Te denken valt aan uitnodigingen voor lezingen, lidmaatschappen van redacties van vooraanstaande tijdschriften, veelvuldige referenties aan het verrichte werk, verworven subsidies en prijzen, eredoctoraten e.d. (Overigens zijn deze 'onderscheidingen' veelal het gevolg van veelvuldige en baanbrekende internationale publikaties op het vakgebied). Ook zal er wel een onderzoeker te bedenken zijn die met

één publikatie het aanzien van zijn vak heeft veranderd (en dus van grote betekenis was voor zijn discipline) en daarna weinig meer heeft laten zien (en dus weinig productief was). De hieronder te presenteren bescheiden citatieanalyse vormt een eerste poging althans één van dergelijke kwaliteitsindices te verdisconten. Wel willen we waarschuwen tegen het met *dédain* neerkijken op het tellen van publikaties en het zonder meer de voorkeur geven aan het eigen kwalitatieve oordeel. Immers, elke internationale publikatie heeft voldaan aan de strenge maatstaven van minstens drie collega's die niet bij voorbaat minder intelligent of erudiet zijn dan wijzelf. Het zou van hovaardigheid getuigen deze oordelen van buitenlandse collega's zonder meer terzijde te schuiven (bijvoorbeeld, omdat gebleken is dat het 'peer-review' systeem niet onfeilbaar is) ten faveure van het eigen diepzinnige oordeel. Met andere woorden, met moet met zeer goede argumenten komen wil men het werk van een internationaal veel publicerende collega op kwalitatieve gronden afwijzen.

Na aldus het belang van internationale publikaties als kwaliteitsaanduiding benadrukt te hebben, gaan we nu over tot de evaluatie van het publikatiegedrag van Nederlandse pedagogen en onderwijskundigen op dit terrein. Hierbij zal blijken dat zelfs tellen enige intelligentie vergt.

Het turven van (inter)nationale publikaties. Enige methodologische problemen

Bij onze berekeningen van de aantallen publikaties zijn we uitgegaan van de gegevens die de instellingen ter beschikking hebben gesteld aan de VSNU-commissie-Knoers. Door elke universiteit is een aantal projecten beschreven en bij elk project is een lijst wetenschappelijke publikaties gegeven voor de jaren 1990-1994 (de zogeheten output). Daarnaast heeft iedere universiteit aangegeven hoeveel onderzoeksuren door elk project in die vijf jaren zijn gespendeerd (de zogeheten input). Gezien de belangen die op het spel stonden zal het niet verbazen dat deze gegevens niet geheel betrouwbaar zijn (zie Kroonenberg & Van der Veer, 1995)¹. We

beperken ons hier tot de rubriceringsproblemen die zich voordoen bij de opgevoerde publikaties.

1) Het eerste probleem rijst bij de definitie van wat mag gelden als een internationaal artikel. Wij hebben ervoor gekozen dit te lezen als 'een publikatie in een internationaal wetenschappelijk tijdschrift'. Dit betekent dat in het Engels etc. gestelde artikelen in Nederlandse tijdschriften door ons naar de categorie Nederlandse publikaties zijn verwezen. Als gevolg hiervan zijn met name enige artikelen in *Come-nius* en het *Tijdschrift voor Onderwijsresearch* als Nederlandse publikatie gerekend.

2) Het tweede probleem is aan te geven wat als artikel mag gelden. We hebben ervoor gekozen (als zodanig herkenbare) boekbesprekingen, 'letters to the editor', en 'abstracts' niet als artikel te erkennen. Uitzonderlijk korte bijdragen (1 à 2 pp.) zijn wel gehandhaafd. Verder is een groot aantal Engelstalige rapporten over (delen van) computerprogramma's, met name bij de RUG, buiten alle tellingen gehouden.

3) Het derde probleem is te bepalen wat een internationaal wetenschappelijk tijdschrift is. Zeer veel van de opgevoerde tijdschriften (bijv., *Insights*; *Linux*; *Rasch Measurement*) kwamen niet voor op de ISI-lijsten, en zijn niet te vinden in het landelijke Online-Picabestand. Slechts in een enkel geval (bijv., *Activity Theory*) was het ons bekend dat het in feite slechts om een nieuwsbrief of clubblad ging. Noodgedwongen zijn we uitzonderlijk liberaal geweest in het accepteren van periodieken als internationaal wetenschappelijk tijdschrift. Het vermoeden bestaat echter dat deze categorie 'vervuild' is door tijdschriften die in feite geen volwassen beoordelingssysteem kennen. Als dit vermoeden juist is, zal deze omstandigheid de relatie tussen internationale produktie en kwaliteit dus doen afnemen.

4) De categorie internationale wetenschappelijke boekpublicaties levert nog veel meer problemen op. Kiest men er voor deze rubriek te definiëren als 'boek(hoofdstukken) bij gerenommeerde internationale of internationaal opererende wetenschappelijke uitgeverij', dan blijkt deze vrijwel leeg. Nogal wat internationale boeken zijn uitgegeven door obscure buitenlandse (bijv., Lisboa: Dom Quixote; Sevilla: Servicio Andaluz de Salud; Moscow: September Publishing Company, etc.) of be-

kende dan wel onbekende Nederlandse uitgevers (bijv., De Lier: ABC; Utrecht: Bishop Bekkers Foundation; Utrecht: Lemma; Amsterdam: Het Spinhuis, Ljouwert: De Fryske Akademy, etc.) van wie het vermoeden bestaat dat zij de bijdragen niet hebben onderworpen aan een 'peer review' proces en dat zij niet in staat zijn boeken in het buitenland te distribueren. In een enkel geval lijkt het internationale boek in feite niet meer te zijn dan een in het Engels gesteld intern of extern rapport, terwijl ook proceedings veelvuldig als boek vermomd blijken te zijn. Bij gebrek aan heldere criteria bleek het voor ons nagenoeg ondoenlijk een scherp onderscheid te maken tussen 'echte' en 'onechte' internationale wetenschappelijke uitgevers, tussen boeken en proceedings, en (soms) tussen boeken en interne rapporten. Noodgedwongen hebben we daarom uiteindelijk alleen het onderscheid 'internationaal artikel vs. internationaal overig' gehanteerd en in de categorie 'internationaal overig' vrijwel alles geaccepteerd. Het vermoeden is dus dat deze laatste categorie in veel ernstiger mate 'vervuild' is dan de categorie 'internationale artikelen';

5) De categorie nationale artikelen leverde geen speciale problemen op. Enige periodieken (bijv., *Documentatieblad Nadere Reformatie*; *Documentatieblad voor de Geschiedenis van de Nederlandse Zending en Overzeese Kerken*) waren ons niet bekend als wetenschappelijk tijdschrift, maar bij gebrek aan heldere criteria hebben we (vrijwel) alles in deze categorie meegeteld.

6) Wat resteert zijn Nederlandse publikaties die niet in een tijdschrift opgenomen zijn (boekpublicaties, handboeken, externe rapporten, proceedings etc.). Als er één categorie is waar rijp en groen op onnavolgbare wijze door elkaar gemengd zijn, dan is het deze. Zeer populair blijken publicaties door GION, RION, ITS, ISOR, SVO en aanverwante instellingen, terwijl ook diverse stichtingen (bijv., Stichting Kinderstudies) vele publicaties hebben uitgegeven. Het komt ons voor dat dergelijke publicaties niet geheel te vergelijken zijn met uitgaven door 'echte' uitgevers. Het schrappen van dergelijke publicaties zou aanzienlijke gevolgen hebben voor sommige projecten (in één geval zouden ruim 80 van de 114 boekpublicaties vervallen).

Op grond van deze overwegingen zou men dus mogen vermoeden dat de vier produktiviteitsmaten niet gelijkwaardig zijn. In het bijzonder valt te verwachten dat tijdschriftpublicaties een betere kwaliteitsindicator vormen dan niet-tijdschriftpublicaties. Anders gezegd, men kan willekeurig welk produkt vermomen als wetenschappelijk rapport of boek, bijvoorbeeld door het zelf uit te geven, maar bij wetenschappelijke tijdschriften ligt dit iets moeilijker. Bij het bepalen van de kwaliteit van een onderzoeksproject kan men dus in het algemeen beter afgaan op het aantal gepubliceerde (internationale) artikelen dan op de overige publicaties. Nederlandse niet-tijdschriftpublicaties vormen een zeer heterogene en sterk vervuilde categorie en een positieve samenhang van het aantal publicaties in deze categorie met de kwaliteit van een onderzoeksproject valt op voorhand niet te verwachten.

De (inter)nationale publicaties in beeld

Allereerst bekijken we de gemiddelde aantallen publicaties per project per fte per jaar (zie Tabel 1). Het aantal internationale tijdschriftpublicaties (volgens onze redenering de beste aanduiding van kwaliteit) blijkt sterk uiteen te lopen van 0.08 tot 2.96 bij een gemiddelde van 0.85. Dat wil zeggen, dat Nederlandse pedagogen (inclusief aio's) gemiddeld bijna één internationaal artikel per jaar publiceren indien zij het gehele jaar aan onderzoek zouden besteden en indien wij alle opgevoerde internationale tijdschriften als zodanig erkennen. Dit gemiddelde is bescheiden, maar ligt overigens niet dramatisch ver onder dat van de psychologen en onder dat wat internationaal gebruikelijk is ('Commissie Frijda', p. 11). Absolute koploper op dit terrein is het project van Duker (2.96), op afstand gevolgd door de projecten van Van der Voort (2.25), Moonen (1.99), Van IJendoorn (1.85), en Van der Linden (1.82). In de categorie 'overige internationale publicaties' (waarvan wij betoogd hebben dat die als kwaliteitsindicator minder betrouwbaar is) vinden we geheel andere koplopers, te weten de projecten van Creemers (3.83), Dronkers (2.83), Van Gent (2.41), Nakken (2.05), en Ruijsseenaars (1.76).

Tabel 1

Overzicht van het gemiddeld aantal publikaties per fte per project (91-94)

Project Univ.	Proj. Leider	Internationaal		Nationaal		ISI ≥ 4	ISI	Geen ISI
		Artikel	Overige	Artikel	Overige			
UT	1 Nijhof	1.49	1.60	0.37	1.08	0.00	0.48	1.00
UT	2 Dijkstra	1.33	1.26	0.31	0.77	0.56	0.87	0.45
UT	3 Moonen	1.99	1.20	0.00	2.10	0.03	0.61	1.37
UT	4 Scheerens	1.16	0.80	1.16	1.56	0.10	0.15	1.00
UT	5 Van der Linden	1.82	1.75	0.29	1.53	0.86	1.39	0.44
RUL	1 Van IJzendoorn	1.85	1.64	1.89	2.60	0.71	1.17	0.68
RUL	2 Van der Voort	2.25	0.56	1.91	2.70	0.45	0.90	1.35
RUL	3 Ruijsenaars	0.97	1.76	2.87	4.63	0.37	0.46	0.51
RUL	4 Van Gent	1.04	2.41	1.74	1.89	0.09	0.09	0.94
RUL	5 Du Bois-Reymond	0.94	1.73	1.22	3.17	0.07	0.14	0.79
RUL	6 Boekaerts	0.95	1.22	0.91	2.13	0.18	0.40	0.55
RUL	7 Ippel	0.08	0.42	0.08	0.50	0.00	0.08	0.00
RUG	1 Dekker	1.55	1.45	1.50	1.61	0.21	0.52	1.04
RUG	2 Nakken	0.48	2.05	1.72	3.92	0.11	0.18	0.29
RUG	3 Van der Kamp	0.42	1.67	1.55	2.74	0.06	0.18	0.24
RUG	4 Koster	0.50	0.93	0.89	1.61	0.04	0.18	0.32
RUG	5 Creemers	0.75	3.83	2.93	8.57	0.15	0.15	0.60
UVA	1 Dronkers	1.63	2.83	3.59	5.76	0.33	0.43	1.20
UVA	2 Van Hout	0.19	1.13	2.26	5.09	0.00	0.00	0.19
UVA	3 De Gloppe	1.28	0.77	2.82	2.82	0.68	0.94	0.34
UVA	4 Van Wieringen	0.14	1.64	0.41	1.74	0.00	0.00	0.14
UVA	5 De Savornin L.	0.56	0.51	0.86	0.81	0.00	0.00	0.56
UVA	6 Hermanns	0.58	1.10	1.33	1.49	0.29	0.42	0.16
UVA	7 Defares	0.22	0.89	1.67	4.44	0.11	0.11	0.11
VU	1 Van Oers	1.16	0.47	1.86	1.74	0.12	0.35	0.81
VU	2 Spiecker	0.79	0.67	1.76	1.76	0.00	0.30	0.48
VU	3 Baartman	0.54	1.02	2.16	4.19	0.00	0.00	0.54
VU	4 Van der Leij	1.13	1.05	1.58	4.77	0.68	0.68	0.45
UU	1 Meeus	0.92	1.01	1.44	2.24	0.34	0.55	0.37
UU	2 Imelman	0.65	1.08	1.08	1.88	0.00	0.00	0.65
UU	3 Kanselaar	0.46	0.84	0.66	3.16	0.02	0.09	0.38
UU	4 Rispens	0.36	0.42	1.19	2.02	0.24	0.24	0.12
UU	5 Stevens	0.18	1.04	0.31	1.84	0.00	0.00	0.18
KUN	1 Gerris	0.54	0.62	0.75	1.78	0.21	0.37	0.17
KUN	2 Van Acker	0.45	0.75	0.45	1.79	0.00	0.00	0.45
KUN	3 De Bruyn	0.42	0.42	0.63	0.91	0.17	0.24	0.17
KUN	4 Simons	0.49	1.20	0.60	1.64	0.08	0.18	0.31
KUN	5 Wildemeersch	0.54	1.01	0.90	2.01	0.00	0.04	0.50
KUN	6 Van Bon	0.71	0.71	1.67	1.61	0.48	0.54	0.18
KUN	7 Duker	2.96	0.85	0.99	1.97	0.99	1.69	1.27
KUN	8 Van Haaften	0.63	0.63	2.02	1.07	0.04	0.08	0.56
Totaal		0.85	1.19	1.25	2.32	0.19	0.35	0.51

Noot: De getallen in de Tabel vertegenwoordigen het gemiddeld aantal publikaties per jaar per fte, waarbij fte is gedefinieerd als het gemiddeld aantal eerste-geldstroom-medewerkers (inclusief aio's) van een eenheid ervan uitgaande dat zij hun volledige tijd aan onderzoek kunnen besteden. In werkelijkheid gaat het dus om twee keer zoveel full-time aanstellingen, uitgaande van onderzoeksinstellingen van 50%.

Legenda: ISI = artikelen uit tijdschriften opgenomen in het ISI-bestand (zie tekst); ISI ≥ 4 = ISI-artikelen met een impact groter gelijk aan 4; Geen ISI = artikelen in internationale tijdschriften die niet opgenomen zijn in het ISI-bestand. De gebruikte impactcijfers zijn die van 1993, of die van 1991 als die van 1993 niet voorhanden waren.

Kijken we naar de nationale tijdschriftpublicaties, dan blijkt dat de Nederlandse pedagogie gemiddeld iets meer dan één artikel (1.25) per jaar in een Nederlands tijdschrift publiceert (weer ervan uitgaande dat het gehele jaar aan onderzoek besteed wordt). De projecten die excelleren op dit terrein, zijn de projecten van Dronkers (3.59), Creemers (2.93), Ruijsse-naars (2.87), De Glopper (2.82), en Van Hout (2.26). In de volgens ons kwalitatief minst betrouwbare categorie 'overige Nederlandse publicaties' vinden wij als topprojecten de projecten van Creemers (8.57), Dronkers (5.76), Van Hout (5.09), Van der Leij (4.77), en Ruijsse-naars (4.63).

Reeds op grond van deze beperkte gegevens zou men kunnen vermoeden dat de categorie 'internationale artikelen' zich onderscheidt van de andere drie categorieën. Immers, geen van de prominenten uit deze categorie blinkt uit in de andere categorieën, terwijl de projecten van Creemers, Dronkers, en Ruijsse-naars gezamenlijk hoog scoren in de categorieën 'overige internationale publicaties', 'Nederlandse artikelen', en 'overige Nederlandse publicaties'. Zou het wellicht zo zijn dat voor de Nederlandse pedagogiek als geheel geldt dat de productie in de laatstgenoemde categorieën sterk samenhangt, terwijl de productie van internationale artikelen een aparte categorie vormt? Teneinde dit na te gaan, hebben we de correlatie tussen de scores in deze categorieën bepaald (zie Tabel 2). Uit Tabel 2 lezen we af dat het aantal internationale artikelen inderdaad niet samenhangt met de productie in de categorie

'Nederlandse artikelen' ($r = .12$) en die in de categorie 'overige Nederlandse artikelen' ($r = -.01$). Verrassender is wellicht dat er ook geen noemenswaardige correlatie is tussen het aantal internationale artikelen en het aantal producten in de categorie 'overige internationale publicaties' ($r = .15$). Deze bevinding, gevoegd bij de relatief hoge onderlinge correlaties (van $r = .41$ tot $r = .70$) tussen de categorieën 'overige internationale publicaties', 'Nederlandse artikelen', en 'overige Nederlandse publicaties', lijkt erop te wijzen dat er inderdaad sprake is van verschillende publikatietradities: er is een groep projecten die bij uitstek publiceert op de nationale markt en in het gebied van de 'overige internationale publicaties' en er is een groep die mikt op het forum van de internationale tijdschriften en die zich soms wel, soms niet in het nationale debat mengt. Gezien onze eerdere argumenten en de gemaakte aantekeningen bij de diverse categorieën denken wij dat de strategie van deze laatste groep wetenschappelijk gezien de voorkeur verdient. De Tabel laat ook zien dat het aantal dissertaties per (niet-aio) eerste geldstroom wp alleen enige samenhang vertoont met de Nederlandse productie. Naar de achtergrond hiervan kunnen we slechts gissen.

Hierboven vermeldde we dat we bij gebrek aan heldere criteria alle opgevoerde internationale tijdschriften als zodanig erkend hebben. We kunnen echter een poging doen een verdere schifting in deze categorie aan te brengen door na te gaan welke tijdschriften opgenomen zijn in de bestanden van de ISI (Institute for Scienti-

Tabel 2
Correlaties tussen Publikatiecategorieën

Publikatiecategorie	Internationaal					Nationaal	
	ISI≥.4	ISI	Geen ISI	Artikel	Overig	Art.	Overig
ISI: impact≥.4	1.00						
ISI	.92	1.00					
Geen ISI	.18	.38	1.00				
Internationale Artikelen	.68	.84	.81	1.00			
Internationaal Overig	.05	.01	.24	.15	1.00		
Nationale Artikelen	.18	.03	.17	.12	.41	1.00	
Nationaal Overig	.06	-.08	.07	-.01	.64	.70	1.00
Dissertaties/fte	.17	.11	-.13	-.01	-.01	.41	.27

Legenda: ISI = artikelen uit tijdschriften opgenomen in het ISI-bestand (zie tekst); ISI≥.4 = ISI-artikelen met een 1993 impact groter gelijk aan .4; Geen ISI = artikelen uit internationale tijdschriften die niet opgenomen zijn in het ISI-bestand. Dissertaties/fte = gemiddeld aantal dissertaties per jaar per gemiddeld aantal eerste-geldstroom medewerkers minus aio's.

fic Information) en welke niet. De ISI selecteert tijdschriften op kwaliteit (onder andere wordt gekeken naar de review-procedure en de mate waarin artikelen geciteerd worden in andere tijdschriften) en beheert thans het meest volledige bestand van internationale tijdschriften dat er bestaat. De kans dat een gezaghebbend tijdschrift in dit bestand ontbreekt is minimaal. Opname van een tijdschrift in dit bestand valt dus als een soort minimale kwaliteitsgarantie te beschouwen. Daarnaast heeft het publiceren in ISI-tijdschriften het voordeel dat een artikel deel uitmaakt van een volledig ontsloten databank (via *Current Contents* en de *(Social) Science Citation Index*), zodat Engelslezende collega's in binnen- en buitenland kennis kunnen nemen van het gerapporteerde onderzoek, zelfs zonder het tijdschrift te bezitten.

Het blijkt dat hantering van het ISI-criterium een kleine verschuiving aanbrengt in de rangorde van de topprojecten op tijdschriftengebied (zie Tabel 1). Het project van Duker (1.69) blijft koploper, terwijl de projecten van Van der Linden (1.39), Van IJzendoorn (1.17), en Van der Voort (0.90) hun positie handhaven. Het project van Moonen (0.61) valt echter weg ten gunste van het project van De Gloppe (0.94). Kijken we naar de correlaties tussen de eerder genoemde categorieën en het aantal publikaties in ISI- dan wel niet-ISI tijdschriften, dan blijkt dat de scores in niet-ISI tijdschriften iets meer samenhangen met de scores in de categorieën 'overige internationale publikaties', 'Nederlandse publikaties', en 'overige Nederlandse publikaties' (zie Tabel 2).

We zouden kunnen proberen de lat nog iets hoger te leggen door zogenaamde impactcijfers in ogenschouw te nemen. Het impactcijfer voor 1994 van, bijvoorbeeld, *Zeitschrift für Pädagogik* wordt bepaald door het aantal verwijzingen in alle tijdschriften uit 1994 naar de in de voorafgaande twee jaar verschenen publikaties in *Zeitschrift für Pädagogik* te delen door het aantal artikelen dat in die jaren verschenen is in *Zeitschrift für Pädagogik*. Impactcijfers zijn problematisch omdat ze mede reflecteren hoe groot de lezerskring van een tijdschrift is. De lezerskring van een tijdschrift wordt echter ook bepaald door toevallige factoren als de omvang van een discipline (psychiatrische tijdschriften hebben, bijvoorbeeld, vaak een hoge

impact) en het al dan niet verbonden zijn van het blad aan het lidmaatschap van een bepaalde beroepsvereniging. Bovendien fluctueren impactcijfers nogal. Idealiter zijn impactcijfers echter geschikt om binnen een bepaalde (sub)discipline na te gaan naar welke tijdschriften het meest verwezen wordt. Dit verwijzingsgedrag hangt waarschijnlijk samen met de status die een tijdschrift heeft. Binnen een bepaalde (sub)discipline kunnen zo frequent geciteerde tijdschriften van tijdschriften met een lagere status onderscheiden worden. In het verleden heeft de 'commissie Frijda' (de commissie die de kwaliteit van het psychologisch onderzoek onderzocht) gekeken naar tijdschriften met een impactcijfer van hoger dan .4 respectievelijk .9. De commissie (Commissie Frijda, 1994, p. 11) was van oordeel dat het laagste criterium het meest voldeed. Als wij dit criterium hanteren (index > .4) dan vinden er wederom geringe verschuivingen plaats in de ranglijst van topprojecten. De drie projecten van Duker (.99), Van der Linden (.86) en Van IJzendoorn (.71) blijven de ranglijst aanvoeren, terwijl het project van Van der Voort (0.45) wegvalt ten gunste van dat van Van der Leij (.68) dat zijn vierde plaats moet delen met het project van De Gloppe (.68).

Dat een auteur zijn artikelen geplaatst weet te krijgen in veel-geciteerde internationale tijdschriften zegt veel over zijn niveau, maar het betekent nog niet dat deze auteur zelf veel geciteerd wordt. Het staat echter buiten kijf dat citaties door vakgenoten een additionele en expliciete erkenning inhouden van de kwaliteit en het belang van iemands onderzoek. Het vaststellen van het aantal malen dat iemand in een zekere periode geciteerd werd in de internationale vakpers is daarom een belangrijke methode om iemands onderzoekskwaliteiten te bepalen. Het is ook een methode die wat verfijnder is dan het mechanische tellen van artikelen en bladzijden en die recht doet aan de hierboven vermelde imaginaire collega die met één hoogwaardige publikatie het aanzien van het vak veranderde. Wij hebben daarom een bescheiden citatie-analyse uitgevoerd voor de periode 1991 tot september 1994 op basis van de SSCI-gegevens. Een dergelijke analyse stelt ons in staat te zien hoe vaak een bepaalde auteur (van een boek of artikel, ongeacht de herkomst) in een ISI-tijdschrift wordt geci-

teerd. Om praktische redenen is deze voorlopige analyse uitgevoerd op eerste auteur, hetgeen een vertekening kan opleveren. Co-auteurs hebben soms een belangrijker aandeel in een publikatie dan een eerste auteur die bijvoorbeeld op een reeks artikelen moet promoveren. In de hoop dat strategieën voor de bepaling van de auteursvolgorde landelijk niet al te sterk verschillen en in de hoop dat eventuele onnauwkeurigheden verzacht worden door aggregatie op projectniveau hebben we niettemin het aantal citaties per project voor de genoemde periode bepaald (zie Kroonenberg & Van der Veer, 1995 voor details). Deze nieuwe kwaliteitsindex levert het volgende beeld. Koploper is het project van Van IJzendoorn (RUL 1) met 135 citaties, gevolgd door de projecten van Simons (KUN 4) met 89, Moonen (UT 3) met 75, Duker (KUN 7) met 38, en Dijkstra (UT 2) met 37. Het is overigens interessant dat aldus bepaalde citatie-cijfers per fte een correlatie laten zien met het aantal publikaties in internationale tijdschriften ($r = .78$ voor internationale artikelen en $r = .75$ voor ISI-tijdschriften).

Conclusies ten aanzien van de produktiviteit

Onze eerste conclusie ten aanzien van de produktiviteit dient een positieve te zijn. Hoewel de getallen die Van IJzendoorn en Van der Veer (1985) vermeldden niet zonder meer vergelijkbaar zijn met de resultaten van onze berekeningen valt toch niet aan de conclusie te ontkomen dat sinds 1985 het aantal internationale publikaties van Nederlandse pedagogen sterk is gestegen. Met tegenzin dan wel *con amore*, aarzelend dan wel trefzeker, de Nederlandse pedagoog is de weg van de internationale publikaties ingeslagen. Kijken we per project naar het aantal internationale tijdschriftpublicaties, dan is het project-Duker (KUN 7) de onbetwiste koploper (ook in 1985 behoorde Duker overigens tot de vier meest producerende pedagogen en onderwijskundigen). Afhankelijk van de gehanteerde meetlat volgen hierop de projecten van De Gloppe (UVA 3), Moonen (UT 3), Van IJzendoorn (RUL 1), Van der Leij (VU 4), Van der Linden (UT 5), en Van der Voort (RUL 2).

Als we naar de produktie van tijdschriftarti-

kelen van universiteiten als geheel kijken, dan valt op dat de Universiteit van Twente (1.57), de Rijksuniversiteit Leiden (1.16) en de Vrije Universiteit (.85) tot de internationaal meest produktieve universiteiten behoren. Zouden we ook de impactcijfers (index > .4) in beschouwing nemen, dan staat de Rijksuniversiteit Leiden (.29) aan de top, gevolgd door de Universiteit van Twente (.25), de Universiteit van Amsterdam (.19) en de Vrije Universiteit (.18).

Als wij toch een kritische opmerking ten aanzien van de internationale produktiviteit zouden moeten maken, dan is het de volgende. Het gebod om internationaal te publiceren wordt nog te vaak opgevat als het gebod om een publikatie in het Engels te stellen. Maar dit zijn twee wezenlijk verschillende zaken. Een internationale publikatie (d.w.z. een publikatie in een echt internationaal tijdschrift of bij een gerenommeerde internationale of internationaal opererende uitgever) ondergaat een strenge intercollegiale toetsing, bereikt veel kritische lezers en voldoet daarmee aan het ideaal dat wij hierboven geformuleerd hebben. In het Engels (Frans, Duits etc.) geschreven brochures, rapporten of boeken die door lokale uitgevers, stichtingen, of vakgroepen worden uitgegeven, worden in het algemeen slechts minimaal getoetst op kwaliteit en bereiken slechts een kleine groep lezers. Dergelijke publikaties zijn slechts in schijn internationaal en dienen onzes inziens op één lijn gesteld te worden met niet-‘gerefereerde’ nationale publikaties. Naar onze mening zijn er nog te veel van dergelijke publikaties en dienen deze publikaties op de internationale markt aangeboden te worden (er van uitgaande dat dit nog niet gebeurd is). Een aantal van deze publikaties zal hierbij sneven omdat zij van lokaal of generlei belang zijn, andere publikaties zullen een verrijking voor de wetenschap blijken. We dienen er dus naar te streven de vele tientallen, zo niet honderden, in het Engels gestelde rapporten, artikelen, boeken, proceedings etc. om te werken tot volwaardige internationale publikaties. Aangezien de barrière om in het Engels te schrijven kennelijk al genomen is, kan deze laatste stap niet groot meer zijn.

Vergelijking met de bevindingen van de 'Commissie Knoers'

Zoals gezegd lijkt het interessant de door ons berekende produktiviteitscijfers te vergelijken met de oordelen die onlangs door de internationale visitatiecommissie onder voorzitterschap van prof. Knoers geveld zijn. Deze commissie heeft met behulp van de standaardcriteria van de VNSU gekeken naar dezelfde gegevens die ten grondslag liggen aan onze berekeningen². De commissie kende cijfers toe aan de kwaliteit, produktiviteit, relevantie, en levensvatbaarheid van het onderzoek die door ons met twee zijn vermenigvuldigd teneinde de overeenkomst met klassieke rapportcijfers (en dus de leesbaarheid) te vergroten.

Een hoog cijfer voor *kwaliteit* kon, onder andere, verkregen worden indien het onderzoek origineel was, een goede bijdrage vormde aan de wetenschap, en indien het niveau van de publikaties hoog was. Bij dit laatste aspect heeft de commissie rekening gehouden met impactcijfers. Helaas wordt niet vermeld hoe dit gebeurd is.

De beoordeling van de *produktiviteit* van een project was, onder andere, afhankelijk van het aantal dissertaties en reguliere publikaties. De commissie maakt gewag van een '1 and 2 rule' die gehanteerd zou zijn ter weging van de 'output' ('Commissie Knoers', p. 7). Helaas is deze regel op volstrekt duistere wijze geformuleerd, zodat ook de oordelen over de produktiviteit van een project niet gereproduceerd kunnen worden³.

De *relevantie* van onderzoek werd bepaald door na te gaan welke bijdrage het onderzoek leverde aan het pedagogisch kennisbestand en

aan de pedagogische praktijk en maatschappelijke ontwikkelingen in het algemeen.

De *levensvatbaarheid* van een onderzoekslijn, tenslotte, hangt volgens de commissie af van zulke praktische zaken als de beschikbaarheid van geld en personeel, de beschikbaarheid van een onderzoeksleider en de interne samenhang van het onderzoek. Ook werd op inhoudelijke gronden een oordeel gegeven over de wenselijkheid van continuering van het onderzoek⁴.

De weinig transparante en daardoor onnavolgbare werkwijze van de commissie en de aperte rekenfouten die zij zich veroorlooft (Kroonenberg & Van der Veer, 1995) geven op voorhand weinig aanleiding de uitgedeelde rapportcijfers al te serieus te nemen. Niettemin zouden de kwalitatieve oordelen van de commissie een bruikbare grove indruk kunnen geven van de stand van zaken in het pedagogisch (en onderwijskundig) onderzoek. In elk geval loont het de moeite de kwalitatieve oordelen van de commissie te vergelijken met onze eenvoudige en volstrekt transparante productiecijfers.

Allereerst zijn we nagegaan wat de correlatie is tussen onze produktiviteitsmaten en de rapportcijfers van de commissie (zie Tabel 3).

Inspectie van Tabel 3 leert ons verschillende dingen.

Ten eerste blijkt er een sterke onderlinge samenhang te bestaan tussen de verschillende oordelen van de commissie (met correlaties variërend van $r = .71$ tot $r = .86$). Er zijn hier twee voor de hand liggende verklaringen mogelijk. Ten eerste is het mogelijk dat deze oordelen sowieso niet volledig onafhankelijk te geven zijn (hetgeen niet onwaarschijnlijk is gezien de omschrijvingen van de oordelen). Ten tweede is het mogelijk dat deze oordelen per project sa-

Tabel 3
Correlaties tussen Publikatiecategorieën en Commissie-oordelen

Beoordelings- Categorie Cie.	Internationaal					Nationaal		Diss. /fte	Categorie		
	ISI	ISI ≥ 4	Art	Geen ISI	Ov.	Art.	Ov.		Kwal.	Lev.	Relev.
Kwaliteit	.52	.48	.42	.16	.02	.05	-.03	.04	1.00		
Levensvatbaarheid	.29	.24	.27	.15	.24	.08	.15	-.04	.82	1.00	
Relevantie	.32	.25	.30	.17	.10	-.03	.03	-.17	.81	.86	1.00
Produktiviteit	.47	.45	.43	.23	.25	.20	.29	-.07	.72	.71	.68

Legenda: ISI = artikelen uit tijdschriften opgenomen in het ISI-bestand (zie tekst); ISI ≥ 4 = ISI-artikelen met een 1993 impact groter gelijk aan .4; Geen ISI = artikelen uit internationale tijdschriften die niet opgenomen zijn in het ISI-bestand. Art. = Artikelen in tijdschriften; Ov. = Overige publikaties. Dissertaties/fte = gemiddeld aantal dissertaties per jaar per gemiddeld aantal eerste-geldstroom medewerkers minus aio's.

menhangen, met andere woorden, dat een goed project automatisch als relevant, productief en levensvatbaar wordt gezien (een zgn. halo-effect)⁵.

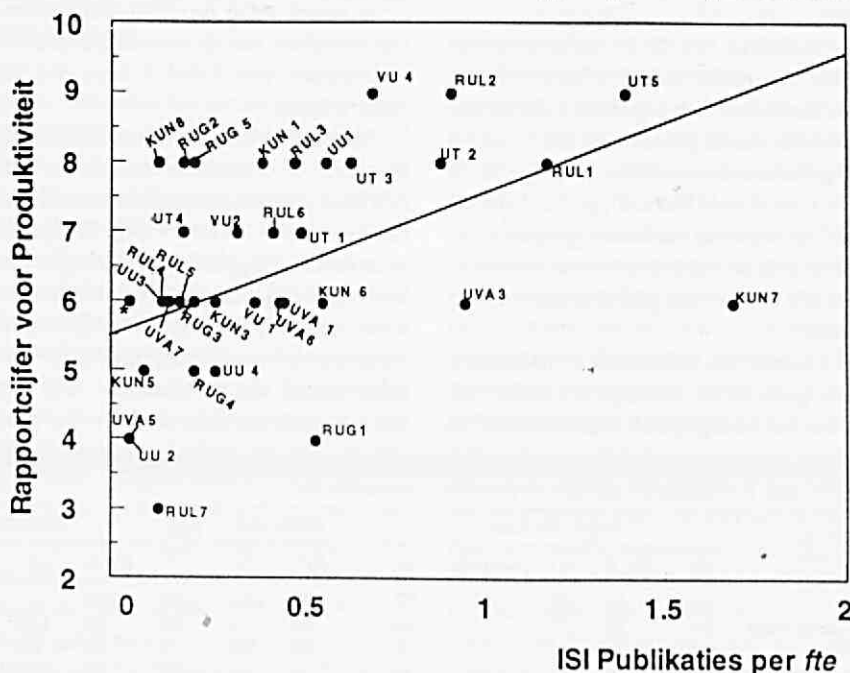
Ten tweede blijkt dat de oordelen van de commissie omtrent kwaliteit, relevantie en levensvatbaarheid niet samenhangen met de door ons berekende aantallen Nederlandse publikaties (de correlatie varieert van $r = -.03$ tot $r = .15$). Kennelijk heeft de commissie bij deze oordelen weinig waarde gehecht aan de objectieve aantallen Nederlandse publikaties. Zelfs het oordeel over de productiviteit correleert niet bijster hoog met de hoeveelheid Nederlandse publikaties ($r = .20$ resp. $r = .29$). Zeer interessant is dat zowel het productiviteitsoordeel als het kwaliteitsoordeel van de commissie betrekkelijk goed correleert (respectievelijk, $r = .47$ en $r = .52$) met de objectieve cijfers voor internationale publikaties in ISI-tijdschriften. Dit betekent dat het aantal publikaties in ISI-tijdschriften de beste objectieve voorspeller is voor de commissie-oordelen en dat de commissie deze publikaties zwaarder heeft gewogen dan andere (internationale) publikaties en indices⁶. Tabel 3 laat ook zien dat het aantal dis-

sertaties per (niet-aio) wp niet correleert met de commissie-oordelen.

Hoewel correlaties in de orde van $r = .50$ interessant zijn, zijn ze natuurlijk niet erg hoog. Dit impliceert dat een en dezelfde productiviteitsscore (als door ons bepaald) gepaard kan gaan met tamelijk uiteenlopende kwaliteits- en productiviteitsoordeelen van de commissie *et vice versa*. Wij kunnen een helder beeld krijgen van de samenhang tussen de commissie-oordelen en de objectieve productiviteitsmaten door per project een regressieanalyse uit te voeren voor de aantallen artikelen in ISI-tijdschriften en, respectievelijk, het productiviteitsoordeel en het kwaliteitsoordeel van de commissie (zie Figuur 1 en Figuur 2).

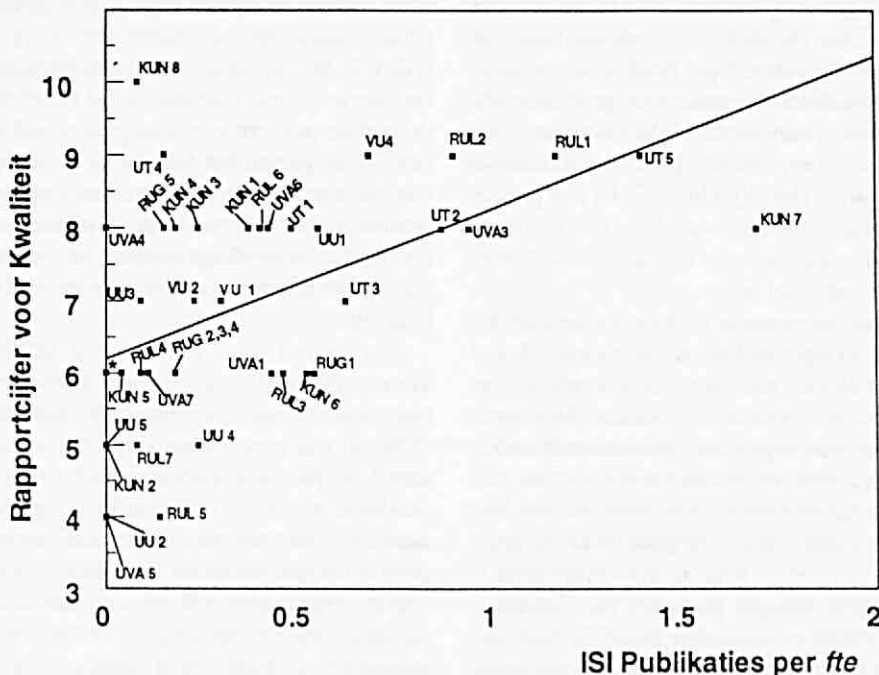
Uit de Figuren 1 en 2 blijken diverse opvallende zaken.

Ten eerste valt op dat een en dezelfde productiviteitsscore inderdaad samen kan gaan met zeer uiteenlopende commissie-oordelen, terwijl anderzijds een en hetzelfde commissie-oordeel zeer verschillende productiviteitscores kan weerspiegelen (zie Figuur 1). Zo zien we, bijvoorbeeld, dat de commissie een zelfde rapportcijfer voor productiviteit, name-



Figuur 1. Rapportcijfers voor de Productiviteit toegekend door de Commissie Knoers afgezet tegen het gemiddeld aantal publikaties per fte per jaar in tijdschriften opgenomen in het bestand van het Institute for Scientific Information (ISI).

Punten zijn gelabeld met de Projectcode (zie Tabel 1); * = KUN2, UVA2, UVA4, VU3.



Figuur 2. Rapportcijfers voor de Kwaliteit toegekend door de Commissie Knoers afgezet tegen het gemiddeld aantal publikaties per fte per jaar in tijdschriften opgenomen in het bestand van het Institute for Scientific Information (ISI).

Punten zijn gelabeld met de Projectcode (zie Tabel 1); * = UVA2 en VU3.

lijk een 6, toekent aan projecten die in objectieve productie in ISI-tijdschriften variëren van .00 tot 1.7. Dit lijkt merkwaardig, omdat een productie van .00 in ISI-tijdschriften beslist laag is (het gemiddelde over alle projecten was .35), terwijl een productie van 1.7 zeer hoog valt te noemen. Hierbij dienen we echter in aanmerking te nemen dat de produktiviteitsmaat van de commissie veel ruimer was dan de onze en, onder andere, ook de Nederlandse publikaties en dissertaties van een project omvatte (zij het dat deze maten weinig correlatie met de produktiviteitsmaat van de commissie vertoonden). Andersom, zien we dat projecten met een objectieve ISI-productie van onder de .10 gewaardeerd worden met rapportcijfers die variëren van 3 tot 8. Dit lijkt eveneens merkwaardig, omdat de commissie zelf grote waarde hecht aan publikatie in internationale tijdschriften en omdat haar produktiviteitsmaat uiteraard deze publikaties omvat. Het lijkt er dus op dat de kwalitatieve produktiviteitsoordelen van de commissie in een aantal gevallen wel erg ver afdwalen van de objectieve produktiviteitscores.

Ten tweede valt op dat bij een aantal projecten de discrepantie tussen de door ons bepaalde produktiviteitsscore en de commissie-oordelen opvallend groot is. Aan de onderzijde van de figuur zien we dat de gemiddeld tot uitstekend producerende projecten van Dekker (RUG 1) en Duker (KUN 7) ver van de regressielijn verwijderd blijven hetgeen op een onderwaardering van de commissie zou kunnen duiden, terwijl aan de bovenzijde van de figuur het project van Van Haaften (KUN 8) een positieve uitschieter vormt.

Bij de relatie tussen het kwaliteitsoordeel van de commissie (Figuur 2) en de ISI-productiecijfers vallen soortgelijke opmerkingen te maken. Weer valt op dat de marges bij een aantal projecten wel erg ruim zijn. De ISI-scores van projecten met het rapportcijfer 8 variëren van 0 tot ongeveer 1.7 (Duker, KUN 7), terwijl projecten met een ISI-score van onder de .25 rapportcijfers voor kwaliteit krijgen die variëren van een 4 tot een 10 (Van Haaften, KUN 8). Absolute uitbijters aan de negatieve kant van de puntenwolk zijn de projecten van Du Bois-Reymond (RUL 5), Rispens (UU 4), en Duker

(KUN 7), terwijl anderzijds het project van Van Haaften (KUN 8) wederom een positieve uitzondering vormt. Opvallend is dat een meer dan gemiddelde internationale produktie een voldoende voorwaarde blijkt voor een hoge kwaliteitsscore, maar geen noodzakelijke voorwaarde. Het blijkt immers dat ook projecten zonder noemenswaardige internationale produktie een goede tot zeer goede kwaliteits-score behaald hebben.

In het voorgaande hebben we gezegd dat men in het algemeen voorzichtig moet zijn met het aan de kant schuiven van de oordelen van internationale tijdschriftredacties. Met name wanneer een project veel internationale artikelen telt, dienen we prudent te zijn met het toekennen van relatief strenge oordelen. Het lijkt daarom gepast eens na te gaan op welke gronden de 'commissie Knoers' in twee gevallen, te weten de Nijmeegse projecten van Van Haaften (KUN 8) en van Duker (KUN 7), haar oordeel zo sterk laat afwijken van de internationale produktiecijfers (zou men deze projecten buiten beschouwing laten, dan stijgt de correlatie tussen het kwaliteitsoordeel van de commissie en het aantal ISI-publikaties van $r = .52$ naar $r = .62$ en die tussen het produktiviteitsoordeel van de commissie en het aantal ISI-publikaties van $r = .47$ naar $r = .62$).

Het internationaal zeer succesvolle project van Duker (KUN 7) krijgt een 6 voor produktiviteit en een 8 voor kwaliteit. De 6 voor produktiviteit wordt gegeven ondanks het feit dat de commissie (p. 75) stelt dat "The group's policy is to publish in international, refereed journals and they have been successful in doing so". De commissie gaat dan echter verder met de woorden "This policy is to be commended, but a more effective method of promoting the group's work within the Netherlands is needed". Kennelijk vindt de commissie op grond van haar uitgangspunten dat de medewerkers van het project van Duker te weinig in het Nederlands publiceren en brengt zij dit oordeel tot uiting in de produktiviteitsscore. De straf voor het uitsluitend internationaal publiceren lijkt echter wel erg hoog uit te vallen. Duidelijk is in ieder geval dat de 6 voor produktiviteit op generlei wijze te rechtvaardigen valt. De kwaliteit van het project van Duker laat in de ogen van de commissie eveneens te wensen over: "the behaviourist approach provides a

weak theoretical framework... The programme needs to elaborate... a coherent theory" (p. 76). Hier lijkt de commissie wel erg hoge maatstaven aan te leggen, maatstaven die kennelijk in de talrijke peer reviews niet gehanteerd werden. Mede gezien het feit dat de pedagogiek van oudsher nogal pluriform van karakter is (waarbij meer 'behavioristische' en andere benaderingen naast elkaar bestaan) had zij er beter aan gedaan zich iets meer op de vlakke te houden.

Het project van Van Haaften (KUN 8) vormt eveneens een geval apart. Hoewel noch het totaal aantal internationale publikaties (1.27 bij een gemiddelde van 2.04) noch het aantal Nederlandse publikaties (3.10 bij een gemiddelde van 3.57) bijzonder hoog valt te noemen, is dit volgens de commissie het enige project dat het predikaat 'excellent' verdient. Opmerkelijk hierbij is dat de commissie meent dat de produktie van het project "may not be adequately reflected in a standard metric", waarbij zij onder meer drie maal verwijst naar een "forthcoming book" dat zij strikt genomen niet in haar oordeel had mogen opnemen. Het komt ons voor dat de commissie zich hier te veel heeft laten leiden door haar subjectief enthousiasme. Een ex aequo plaatsing van dit project met enige andere projecten was onzes inziens meer op zijn plaats geweest.

Conclusies

De Nederlandse pedagogen en onderwijskundigen zijn de afgelopen 10 jaar meer gaan publiceren en mikken in toenemende mate op de internationale markt. Het gemiddelde aantal internationale artikelen is bescheiden, maar ligt niet ver beneden de internationaal gebruikelijke gemiddelden. Overigens zijn de cijfers enigszins geflatteerd, omdat geen poging is gedaan een schifting aan te brengen in de opgevoerde internationale tijdschriften. Wat een acceptabel gemiddelde is voor personen of projecten, is ter beoordeling aan de wetenschappelijke gemeenschap. Voor personen valt te denken aan een norm zoals die geldt voor het lidmaatschap van de ISED-onderzoeksschool: per senior-onderzoeker 5 internationale publikaties per 5 jaar, waarvan 3 in een tijdschrift met een volwassen peer-review systeem. Voor

projecten is het lastiger een goede norm te verzinnen, maar ten opzichte van de door de 'commissie Knoers' gehanteerde '1 and 2 rule' boekt men al gauw enige winst.

Helaas wordt nog onevenredig veel energie gestoken in het produceren van Engelstalige rapporten, boeken e.d. die de internationale markt nooit bereiken. De stelling dat deze rapporten e.d. eigenlijk aangemerkt kunnen worden als een soort verkapte Nederlandse publicaties vindt bevestiging in de gevonden correlaties.

Er blijkt een bescheiden aantal projecten te zijn dat excelleert op de internationale markt. Deze onderzoekers publiceren ook in ISI-tijdschriften van gemiddelde en meer dan gemiddelde impact. Uit hun voorbeeld blijkt dat het mogelijk is tot de internationaal vooraanstaande tijdschriften door te dringen. Overigens dienen Nederlandse pedagogen in voorkomende gevallen tijdschriften voor te dragen voor opname in het ISI-bestand. Voor binnenlands gebruik kan dit bestand (tijdelijk) aangevuld worden met ontbrekende goede tijdschriften.

Onze bevindingen sporen slechts ten dele met die van de 'commissie Knoers'. Het is onduidelijk waarom de kwalitatieve produktiviteitsoordelen van de commissie vaak zover afwijken van de door ons bepaalde objectieve produktiviteitscijfers. Een dergelijke afwijking is onzes inziens niet acceptabel. Toekomstige visitatiecommissies dienen duidelijker aan te geven hoe zij tot hun produktiviteitsoordeel zijn gekomen.

De door de 'commissie Knoers' gegeven kwaliteitsoordelen wijken in een aantal gevallen te ver af van de oordelen die gewettigd zouden zijn op grond van de internationale produktiecijfers. Waar men er van uit mag gaan dat publikaties in internationale gezaghebbende tijdschriften een zware toetsing ondergaan, geeft het onzes inziens geen pas op dit terrein buitengewoon succesvolle collega's een relatief lage beoordeling te geven. Onze cijfers geven ook weinig ondersteuning voor de rangorde van de universiteiten die de 'commissie Knoers' in haar rapport presenteert ('Commissie Knoers', pp. 27-29).

De aan de 'commissie Knoers' geleverde gegevens zijn niet geheel betrouwbaar. De VSNU dient voor toekomstige visitaties strak-

kere richtlijnen voor de te leveren gegevens te hanteren. Aldus gefilterde gegevens kunnen dienen voor objectieve en transparante berekeningen als de onze en dienen een belangrijk element te vormen in de beoordeling van de kwaliteit van onderzoeksprojecten.

Noten

1. Nauwkeurige bestudering van de door de instellingen aangeleverde gegevens leert, bijvoorbeeld, dat de definities van wat geteld wordt als onderzoeksuren niet altijd dezelfde zijn. Ook verschillen projecten wat betreft het aantal aio's in het eerste-geldstroom gedeelte. Het komt ons voor dat de VSNU door strakkere richtlijnen voor de input- en outputgegevens veel problemen had kunnen voorkomen.
2. Helaas blijken de door ons gevonden cijfers nogal wat af te wijken van die van de commissie Knoers. Een analyse van mogelijke oorzaken is te vinden in Kroonenberg en Van der Veer (1995).
3. De meest waarschijnlijke interpretatie van deze regel is dat een project per fte per jaar minimaal 1 internationaal 'gerefereerd' artikel en minimaal 2 niet-'gerefereerde' internationale artikelen dient te hebben. Een dergelijke regel is echter onzinnig, omdat zij een project met uitsluitend internationale 'gerefereerde' artikelen (de hoogste categorie) zou benadelen. Er lijkt dan ook geen enkele reden de projecten van Creemers en Dronkers naar voren te halen als internationale exponenten van de Nederlandse pedagogiek en onderwijskunde (zoals gebeurt op p. 10 van het rapport van de commissie-Knoers en in het persbericht bij het verschijnen van dit rapport) omdat deze projecten als enige aan een dergelijke ongerijmde regel voldoen.
4. Men kan zich afvragen in hoeverre het kwaliteitsoordeel van de commissie gezien kan worden als het resultaat van de andere 3 indices. Een regressie-analyse van Kwaliteit op Produktie, Relevantie, en Levensvatbaarheid levert een $R = .86$ op, hetgeen maar weinig hoger is dan de correlatie tussen Kwaliteit en Levensvatbaarheid ($r = .82$) alleen.
5. Dat het mogelijk is deze oordelen meer onafhankelijk te geven, blijkt uit de gegevens van de commissie-Frijda. De correlaties tussen de oordelen lagen tussen de $r = .39$ en de $.82$ bij een gemiddelde correlatie van ongeveer $r = .60$.

6. Interessant is dat na ISI-publikaties de categorie 'overige Nederlandse publikaties' de enige serieuze predictor ($r = .29$) is voor het produktieoordeel van de commissie ($R = .57$). Het aantal ISI-publikaties blijkt verder de enige serieuze predictor voor kwaliteit te zijn. Alle 'onafhankelijke' variabelen samen komen niet verder dan $R = .53$, dat wil zeggen .01 hoger dan ISI-publikaties alleen.

Literatuur

- Beekman, T. (1986). De internationale status van de Nederlandse pedagogiek. Een brandende kwestie? *Pedagogische Verhandelungen*, 9, 65-68.
- Commissie Frijda (1994). *Quality Assessment of Research. Psychological Research in the Netherlands: Past Performance and Future Perspectives*. Utrecht: VSNU.
- Commissie Knoers (1995). *Quality Assessment of Research. Pedagogical and Educational Sciences in the Netherlands: Past Performance and Future Perspectives*. Utrecht: VSNU.
- IJzendoorn, M. H. van, & Veer, R. van der (1985). De internationale status van de pedagogiek. *Pedagogische Verhandelungen*, 8, 4-12.
- IJzendoorn, M. H. van, & Veer, R. van der (1986). Beekmans blijde boodschap. *Pedagogische Verhandelungen*, 9, 69.
- Kroonenberg, P. M., & Veer, R. van der (1995). *Sluwe pedagogen en verwarde rekenmeesters. Commentaar op het rapport 'Pedagogical and Educational Sciences in the Netherlands: Past Performance and Future Perspectives'*. Leiden: Intern Rapport.
- Merton, R. K. (1973). *The Sociology of Science. Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Manuscript aanvaard 21-3-1996

Auteurs

P. M. Kroonenberg is als Universitair Hoofddocent verbonden aan de Vakgroep Algemene Pedagogiek van de Rijksuniversiteit Leiden.

R. van der Veer is als Universitair Hoofddocent verbonden aan de Vakgroep Algemene Pedagogiek van de Rijksuniversiteit Leiden.

Adres: Vakgroep Algemene Pedagogiek, RUL, Postbus 9555, 2300 RB Leiden

Abstract

International publications and research quality.

An investigation into the publication behavior of Dutch pedagogues and educational scientists

P. M. Kroonenberg & R. van der Veer. *Pedagogische Studiën*, 1996, 73, 225-238.

It is argued that the number of publications in international scientific journals reflects the quality of research projects. An assessment of the (inter)national publication policies of Dutch pedagogues and educational scientists is presented using data supplied by the universities. The conclusions of the International Review Committee of Pedagogical and Educational Sciences are critically discussed in the light of the findings.

De onderzoeksvisitatie 1995

Over visitaties is al heel wat geschreven en gediscussieerd. Wij pretenderen niet over dat onderwerp het laatste, liefst verlossende, woord te kunnen schrijven. We beperken ons tot enkele relativerende opmerkingen.

Allereerst stellen wij vast dat het alleszins nuttig is dat op gezette tijden evaluaties van onderzoeksprogramma's plaatsvinden. Het kan leiden tot (zelf-)reflectie, het kan aanleiding zijn voor faculteits- en universiteitsbesturen het beleid bij te stellen, het kan impulsen geven voor het optimaliseren van onderzoeksactiviteiten op een bepaald onderzoeksgebied. Evaluaties worden tegenwoordig verricht door zogenaamde visitatiecommissies. Men zou natuurlijk wel moeten eisen dat zulke commissies zo objectief mogelijk te werk gaan. Objectiviteit en herhaalbaarheid dienen zo hoog in het vaandel te staan, dat de gepubliceerde evaluatie aan elementaire eisen van wetenschappelijkheid voldoet.

In de gebruikelijke evaluatiepraktijk, zo ook bij visitaties, vindt de beoordeling plaats op programma-niveau. Dat roept meteen allerlei vragen op. Zijn de, gewoonlijk verzwezen, vooronderstellingen over de programma-gewijze organisatie van onderzoek wel valide? Is het bijvoorbeeld inderdaad zo dat de nationale wetenschap alleen vooruit te brengen is door lokale onderzoeksgroepen? Zou niet het alternatief van individuele projecten een levensvatbare mogelijkheid kunnen zijn? Of het alternatief van landelijk georganiseerde inter-universitaire projectgroepen? En als men groepen vergelijkt, zou dan niet allerlei lokaal beleid verdisconteerd moeten worden? Zoals de beschikbaarheid van a.i.o. plaatsen, postdoc posities, infrastructurele voorzieningen, etc.?

Het definiëren van evaluatiecriteria dient te berusten op welomschreven overwegingen met betrekking tot wenselijk wetenschapsbeleid. Visitatiecommissies zouden bij uitstek deskundig moeten zijn in het helder formuleren van degelijke beleidsoverwegingen, die op hun beurt geoperationaliseerd worden in eenduidige beoordelingscriteria. Alweer: de beoorde-

lingen dienen te voldoen aan elementaire wetenschappelijke criteria.

We tekenen aan dat bij het definiëren van evaluatiecriteria aandacht moet worden besteed aan discipline-specifieke criteria. Disciplines verschillen bijvoorbeeld in de mate waarin lokale kennisproductie afhankelijk is van internationale samenwerking tussen wetenschappers. Voor een universitaire onderzoeksgroep op het terrein van de kernfysica is internationale samenwerking een noodzaak; de beoordeling dient daar ook op afgestemd te zijn. Een onderzoeksprogramma Nederlands Recht daarentegen is in haar kennisproductie minder afhankelijk van internationale contacten, beoordelaars dienen daarmee rekening te houden. Soortgelijke overwegingen gelden voor de beoordeling van het succes waarmee derde geldstroomonderzoek wordt aangetrokken. In de biochemie is de productie van fundamentele kennis voor een belangrijk deel afhankelijk van extern gefinancierde projecten, onder meer door het bedrijfsleven. Een visitatiecommissie biochemie zou de desbetreffende programma's derhalve ook op de werfkracht inzake externe gelden moeten beoordelen. Voor onderzoeksprogramma's klassieke talen is een dergelijk criterium niet of nauwelijks van belang.

Een zeer belangrijk punt betreft het doel van evaluaties. Constateren dat een bepaald programma internationaal onbekend is en qua publicaties vooralsnog slechts lokale communicatie heeft bewerkstelligd, is iets gans anders dan vaststellen dat een programma 'nauwelijks levensvatbaar' is. In het eerste geval zegt men dat een programma zich onvoldoende profileert voor het Forum van de wetenschap. In het tweede geval suggereert men dat een programma beter beëindigd kan worden. In de gebruikelijke visitaties is het thans heel gewoon dat uitspraken van de tweede soort gedaan worden. Aangezien er geen reden is om aan te nemen dat visitatiecommissies zouden moeten willen optreden als verkapte bezuinigingscommissies, zou men bij voorkeur willen dat verregaande uitspraken over bijvoorbeeld 'levensvatbaarheid' vergezeld worden van suggesties ter verbetering. In die suggesties zou vooral de zin van de evaluatie kunnen steken.

Als we het rapport van de visitatiecommissie-Knoers bezien vanuit de zojuist ge-

noemde beginselen, dan is er wel wat kritisch commentaar mogelijk. Op vragen als: waarom is het aggregatieniveau van het programma gekozen?; op welke overwegingen met betrekking tot wetenschapsbeleid berusten de gehanteerde beoordelingscriteria?; hoe transparant zijn deze criteria, en welk doel wordt met de evaluatie beoogd?; en zijn zinvolle en realistische suggesties voor de toekomst mogelijk?; ontbreken de antwoorden. Ook heeft de commissie-Knoers zich niet gewaagd aan een positiebepaling van de disciplines pedagogiek en onderwijskunde en de daaruit voortvloeiende discipline-specifieke criteria (terloops, bij de bespreking van orthopedagogiek zijn wat noties te vinden; toekomstige commissies zullen op dit punt hopelijk principiële en uitgebreider ingaan).

Het is de verdienste van de analyses van Kroonenberg en Van der Veer (in dit nummer) dat ze op zakelijke en heldere wijze een aantal discutabele aspecten van het rapport van de commissie-Knoers zichtbaar maken. Uit hun zorgvuldige analyses volgt niet dat het vele en intensieve werk dat door de commissie verricht is geen waarde heeft, maar wel dat op belangrijke punten relativeringen gewenst zijn. Een van de belangrijkste uitkomsten van Kroonenberg en Van der Veer is de gebrekkige samenhang tussen de uitkomsten van de visitatiecommissie en hun eigen, op objectieve criteria gebaseerde evaluaties. Dit is een tot relativering manend gegeven – temeer waar de commissie aangeeft een groot gewicht te hechten aan dezelfde criteria die Kroonenberg en van der Veer toepassen. Als men de visitatiecommissie een zeker recht op subjectieve voor- en afkeuren zou willen toestaan, dan dienen die ons inziens ook expliciet als zodanig te boek te staan. Om ons nu maar te beperken tot het meest extreme voorbeeld: dat zeer lage objectieve scores voor internationaal publiceren kunnen samengaan met kwaliteitscijfers variërend van 4 tot 10 is op zijn minst merkwaardig en vereist dus een heldere uitleg, die in het huidige rapport van de visitatiecommissie geheel ontbreekt. Wat de transparantie van de criteria betreft: deze is niet steeds even optimaal. We volstaan hier met te verwijzen naar de zogenaamde '1 and 2 rule', en daarover door Kroonenberg en Van der Veer gemaakte opmerkingen. Ook op het punt van het doen van

suggesties ter verbetering is de commissie niet altijd even royaal. Mogen wij alweer volstaan met slechts één voorbeeld: het UU2-programma. Dit programma wordt door de commissie nauwelijks levensvatbaar geacht, maar suggesties om de toekomstmogelijkheden te verbeteren ontbreken. Dit voorbeeld maakt tevens duidelijk dat de commissie ook geen rekening lijkt te houden met variaties in de levensloop van een programma. Moeten recent gestarte programma's met een geheel nieuw personeelsbestand gemeten worden met dezelfde maatstaven als reeds langlopende programma's met een vaste personele kern? Moet opstarten van nieuwe programma's worden ontmoedigd?

De voorgaande opmerkingen zijn niet bedoeld om de commissie-Knoers te desavoueren. Men mag verheugd zijn dat een aantal gerenommeerde collega's hun tijd en energie hebben willen besteden aan dit zo moeilijke beoordelingswerk. Onze opmerkingen willen alleen onderstrepen dat beseft dient te worden dat ook dit beoordelingswerk een meta-analyse behoeft (en verdient) en dat deze eerste beoordeling van de pedagogiek en de onderwijskunde feilbaar is. Het ware te wensen dat intercollegiaal en interuniversitair uit het rapport van de visitatiecommissie niet zozeer vergelijkende conclusies worden getrokken (zoals Kroonenberg en Van der Veer opmerken: het gaat hier niet om een sportwedstrijd; wij bevelen aan Kroonenberg en Van der Veer's artikel evenmin voor dat doel te gebruiken), maar vooral wordt nagedacht over beleidsmaatregelen met betrekking tot de toekomst. Maar allereerst is het van belang dat wordt nagedacht over beleidsuitgangspunten, die zinvol zijn voor het onderhavige vakgebied. Toekomstige visitatiecommissies kunnen daarmee hun voordeel doen en de VSNU zou duidelijker dan nu het geval is kunnen aangeven vanuit welk perspectief men het gepubliceerde rapport dient te lezen. De vergelijkende conclusies, die nu met stelligheid in de media zijn terechtgekomen zijn niet gerechtvaardigd. Onze compassie gaat dan ook vooral uit naar die collega's, wier onderzoeksprogramma openlijk is voorzien van treurige rapportcijfers, zonder dat deze cijfers voldoende gerechtvaardigd zijn.

Ten slotte willen wij, om misverstand te voorkomen, nog opmerken dat wij met de voor-

gaande opmerkingen niet willen suggereren dat de door Kroonenberg en Van der Veer gesuggereerde criteria zonder meer moeten worden overgenomen. Natuurlijk: internationaal publiceren is van het grootste belang en geeft een garantie van door deskundige vakgenoten erkende kwaliteit. In menig opzicht kan men aan de door de auteurs gegeven analyses zien dat op logische, zowel als empirische, gronden vooralsnog het door de auteurs voorgestelde internationale publikatie-criterium het best voldoet. Desalniettemin is de exclusieve toepassing van een dergelijk criterium natuurlijk wel degelijk voor discussie vatbaar. Wij noemen de volgende problemen. Moeten wij het als het hoogste goed beschouwen dat de Nederlandse onderzoeker er in slaagt zijn druppeltje aan de oceaan van internationale publikaties toe te voegen? Zou de Nederlandse belastingbetaler tevreden zijn met een situatie waarin wetenschappelijke artikelen over pedagogisch en onderwijskundig onderzoek alleen in internationale tijdschriften verschijnen? Op termijn kan een eenzijdige accentuering op internationaal publiceren leiden tot een waterscheiding: wetenschappelijke artikelen in internationale tijdschriften voor een wetenschappelijk forum en populariserende artikelen in lokale tijdschriften voor professionals en geïnteresseerde leken. Is een dergelijke scheiding wenselijk voor disciplines als de pedagogiek en de onderwijskunde die hun legitimering in belangrijke mate ontleen aan hun toepassingsgerichte opdracht? Zijn het niet juist de professionals die de vertaalslag moeten maken van resultaten

van wetenschappelijk onderzoek naar de praktijk en derhalve ook gemakkelijk toegang moeten hebben tot de wetenschappelijke kennisproductie? Of zijn wetenschappelijke onderzoekers van mening dat onze toevoeging van 'meer van hetzelfde' aan de internationale literatuur het hoogste goed is? Moet uit de analyses van Kroonenberg en van der Veer worden afgeleid dat met het criterium 'internationale artikelen' *niet* samenhangende produktie-indices niet van belang zijn? Als het antwoord op die vraag ja zou zijn, dan volgt daaruit bijvoorbeeld dat dissertaties geen zin hebben (de hoeveelheid dissertaties blijkt volgens de data van Kroonenberg en Van der Veer met geen enkele andere index te correleren!), evenmin overigens als bijvoorbeeld onderwijsresearch-dagen of andere conferenties.

Onze voorgaande opmerkingen zijn bedoeld om de lezer te stimuleren tot nadenken over wetenschapsbeleid voor pedagogiek en onderwijskunde, over produktie- en kwaliteitscriteria, en over vastatiecommissies. En wij hopen tevens dat de lezer zich er meer en beter van bewust is dat de gerapporteerde evaluaties met grote terughoudendheid gehanteerd dienen te worden. Laten we hopen dat de VSNU bereid is lering te trekken uit de reacties op deze eerste evaluatie van de pedagogiek en de onderwijskunde en dat de Universiteits- en Faculteitsbesturen hun beleid niet op ongefundeerde wijze zullen laten beïnvloeden door de gepubliceerde evaluaties.

Redactie Pedagogische Studiën

Boekbesprekingen

J. Peschar & A. Wesselingh

Onderwijssociologie

Wolters-Noordhoff, Groningen 1995 294 pagina's
f 59,50 ISBN 90 01 70691 6

In het afgelopen jaar heeft Wolters-Noordhoff vrijwel al haar inleidingen in de onderwijswetenschappen geheel vernieuwd. Ook de oude inleiding in de onderwijssociologie uit 1985 onder redactie van Peschar en Wesselingh is vervangen door een nieuw boek met een andere opzet.

Het eerste dat opvalt is dat deze inleiding geheel door Peschar en Wesselingh geschreven is en niet door twintig verschillende auteurs. Deze aanpak is zowel de omvang (deze is bijna gehalveerd) als de inhoud ten goede gekomen. De vorige inleiding miste elke structuur en bevatte ook hoofdstukken die slechts met veel moeite sociologisch genoemd konden worden (bijvoorbeeld de hoofdstukken over economie en recht). De nieuwe opzet heeft uiteraard ook zijn grenzen. Twee auteurs kunnen een minder breed veld bestrijken dan twintig specialisten.

De systematiek van de nieuwe inleiding wordt gevormd door twee leidraden: de hoofdvragen van de sociologie en de onderzoekscyclus. Beide zijn ontleend aan het leerboek *Sociologie* van Ultee, Arts en Flap (1992). Volgens Ultee e.a. zijn er drie centrale, sociologische vraagstukken: ongelijkheid, cohesie en rationalisering. Deze hoofdvragen zijn, zoals ook uit diverse besprekingen van Ultee e.a. is gebleken, overigens aanvechtbaar als dekking voor de gehele sociologie. Desalniettemin hebben Peschar en Wesselingh deze vertaald naar de onderwijssociologie in het vraagstuk van te verwerven competenties en kwalificaties, het vraagstuk van ongelijkheid en differentiatie, en het vraagstuk van integratie in en door het onderwijs.

In het tweede deel van het boek worden deze vraagstukken steeds in drie hoofdstukken behandeld. Binnen een hoofdstuk wordt de onderzoekscyclus als leidraad gehanteerd. Tot de onderzoekscyclus rekenen Ultee e.a. drie stappen: eerst wordt er een *probleem* gesteld, dan wordt er een *theorie* gezocht of geformuleerd en daarna wordt er *onderzoek* gedaan. Peschar en Wesselingh breiden deze cyclus met nog een stap uit: het *beleid*. Daarvoor geven zij twee motieven. In

de eerste plaats zou veel onderwijssociologisch onderzoek tot gevolg hebben dat er specifiek beleid bedacht of uitgevoerd wordt. Dat lijkt mij wat erg veel eer voor de onderwijssociologie. Studies naar de rol van onderzoek in het beleid wijzen daar in ieder geval niet op. Wat niet wegneemt dat betrekkelijk veel onderwijssociologen een rol in beleid spelen of gespeeld hebben. Dat is echter heel wat anders. Het tweede motief is dat verschillende problemen in de onderwijssociologie rechtstreeks gerelateerd zijn aan beleidsvragen. Ook dat lijkt mij geen sterk motief. Veel vooraanstaande sociologen (bijvoorbeeld Bourdieu) hebben juist gewezen op het gevaar van een te sterke relatie tussen beleidsproblemen en wetenschappelijke problemen.

In het eerste deel van het boek geven de auteurs in vier hoofdstukken een algemene inleiding. Hoofdstuk 1 begint met een brede bespreking van het sociologisch gezichtspunt, dat in het daarop volgende hoofdstuk wordt toegespitst op de sociologie van het onderwijs. In het derde hoofdstuk gaan de auteurs in op het ontstaan en de ontwikkeling van de Nederlandse onderwijssociologie en ten slotte schenken zij in het vierde hoofdstuk aandacht aan de theorie en het onderzoek van de onderwijssociologie. Deze vier hoofdstukken vertonen nog de meeste gelijkenis met de drie inleidende hoofdstukken uit het vorige boek.

De (herziene) geschiedenis van de Nederlandse onderwijssociologie is een aardig hoofdstuk geworden. Het is echter duidelijk geschreven door een betrokkene en dat weerkent zich bijvoorbeeld in de paragraaf 'verschillen tussen scholen en hun organisatie'. Veel interessante conceptualisering en studies ontbreken, maar ook de indeling is wat merkwaardig. Organisatiesociologische benaderingen van het onderwijs (waarbij een van de belangrijkste stromingen ontbreekt: het neo-institutionalisme) worden bijvoorbeeld op een geforceerde wijze ondergebracht bij het effectiviteitsonderzoek. Interessant zijn enkele passages in de afsluiting van het hoofdstuk. Daar wordt feitelijk een paradox geconstateerd: enerzijds is de onderwijssociologie door verlies aan vertrouwen "niet geheel onbeschadigd uit de crisis binnen de sociale wetenschappen" gekomen, anderzijds heeft de onderwijssociologie "de moeilijkheden goed doorstaan, vooral dankzij haar solide empirische basis." Anders gezegd: ondanks meer en beter

onderzoek, is er minder (publiek) vertrouwen. Het zou interessant geweest zijn als de auteurs op een of andere manier daar op waren ingegaan, bijvoorbeeld vanuit een meer wetenschapssociologisch gezichtspunt.

De drie andere inleidende hoofdstukken zijn vooral bedoeld voor niet-sociologen om kennis te maken met de sociologische visie op maatschappelijke verschijnselen als het onderwijs. Hoewel heel helder geschreven, is dat niet altijd geslaagd. Zo komen in het eerste hoofdstuk de drie hoofdvragen van de sociologie (ongelijkheid, cohesie en rationalisering) geheel uit de lucht vallen. In de paragrafen daarvoor zijn namelijk geheel andere centrale begrippen en theoretische stromingen de revue gepasseerd. Dit is niet alleen jammer voor niet-sociologen, maar ook voor vakgenoten die de verdere indeling van het boek willen begrijpen. Dergelijke grote sprongen zijn ook in het vierde hoofdstuk te vinden. Na een schets van de hoofdrichtingen binnen de onderwijssociologie volgt een paragraaf waarin twee willekeurige onderzoeken worden besproken, met als conclusie dat de keuze voor een onderzoeksmethode afhankelijk is van de vraagstelling. Die twee onderzoeken zijn daar geen illustratie van. Het zou beter geweest zijn om nu juist die afhankelijkheid met enkele voorbeelden te laten zien. In plaats daarvan volgt nu een uitvoerige paragraaf over school- en beroepsloopbanenonderzoek. Hoe interessant deze paragraaf op zich ook is, zij toont weinig van de relatie tussen theorie en empirie.

Het tweede deel bevat een aantal interessante hoofdstukken dat zeker zijn nut in het onderwijs zal bewijzen. Vooral de hoofdstukken waaruit de deskundigheid van de auteurs blijkt, zijn niet alleen lezenswaardig, maar ook behoorlijk degelijk. Zo geven de hoofdstukken over de aansluiting van onderwijs en arbeid, onderwijs en maatschappelijke ongelijkheid en onderwijs als stijgingskanaal een goed inzicht in de stand van de Nederlandse onderwijssociologie. Maar ook de meer theoretisch getinte hoofdstukken over de plurale samenleving en normen en waarden zijn leuke overzichtsstukken, die een goed uitgangspunt voor verdere discussie en studie vormen.

Desalniettemin wrekt zich in veel hoofdstukken het eerder genoemde PTOB-schema (probleem, theorie, onderzoek en beleid) dat de auteurs zich opgelegd hebben. Niet elk onder-

werp leent zich even goed voor dit schema. Dit schijnt de auteurs ook te storen, want soms lopen de stappen flink door elkaar: stukken theorie worden al onder het kopje probleem behandeld; onder het kopje beleid wordt ook onderzoek behandeld; onder beleid worden nu eens nota's en wetten uitvoerig besproken, dan weer wat globale beleidsproblemen kort aangeduid; ten slotte wordt in enkele gevallen een afwijkend schema gehanteerd.

Als het hier alleen om een stijlprobleem zou gaan, zouden deze inconsistenties niet zo bezwaarlijk zijn. Er gaat echter een veel fundamenteeler probleem achter schuil. Het schema veronderstelt een zekere rationaliteit en lineariteit die in de werkelijkheid niet bestaat. In de eerste plaats blijft vaak onduidelijk of de auteurs onder een probleem nu een maatschappelijk of een wetenschappelijk probleem verstaan. Dit varieert niet alleen per hoofdstuk, maar ook binnen de hoofdstukken. In de tweede plaats geven de auteurs weinig rekenschap van het feit dat wetenschappelijke problemen theoriegeladen zijn. Uit een bepaalde probleemdefiniëring spreekt een zekere theoretische visie. Daardoor ontkomt men er niet aan om in de probleemparagraaf al theoretische inzichten te bespreken. In de derde plaats zien de auteurs klaarblijkelijk te weinig in, dat een beleidsprobleem van een andere orde is dan een wetenschappelijk en ook sociaal probleem. Bij beleidsproblemen spelen processen van agendavorming (wie plaats wat, op welke wijze, op de politieke agenda) en ideologievorming (de waardenbeïnvloeding van de probleemdefinitie) een grote rol.

Een voorbeeld van een hoofdstuk waar dit fundamentele probleem niet alleen de opbouw, maar ook de inhoud verstoort is hoofdstuk 12 over verzuiling, groepsidentiteit en emancipatie. Het hoofdstuk begint met een korte schets van het ontstaan van het verzuilde onderwijs. Volgens de auteurs zou het verzuilde stelsel vooral het resultaat zijn van de instelling van het Koninkrijk der Nederlanden in 1814 en de emancipatie van de Nederlandse rooms-katholieken sinds het herstel van de bisschoppelijk hiërarchie in 1853. Merkwaardig in deze schets is dat de rol van de protestanten in het geheel niet genoemd wordt. Ook de rol van de socialisten wordt op een wat vreemde (zelfs anachronistische) manier getypeerd.

Daarna volgt een paragraaf over verzuiling

naar levensbeschouwelijke diversiteit. Nog afgezien van de rommelige opbouw spreken de auteurs zichzelf op een nogal belangrijk punt tegen. De auteurs omschrijven *verzuiling* als het bestaan van min of meer afgesloten groepsverbanden die op basis van levensbeschouwing gericht zijn op het vervullen van uiteenlopende maatschappelijke functies op diverse maatschappelijke terreinen. Het begrip *verzuiling* lijkt mij in dit verband wat ongelukkig gekozen, omdat dat zou moeten verwijzen naar een proces en niet naar een bestaanswijze. Vervolgens komen de auteurs met de vaststelling dat *verzuiling* geen exclusief Nederlands verschijnsel is. Dit is een verdedigbare stelling, maar niet op grond van de door hen gehanteerde definitie! Dat blijkt ook uit de voorbeelden die zij geven: taalgroepen, regionale en lokale verschillen. Vervolgens stellen zij dat wel uniek voor ons land is, dat religie de grondslag van de *verzuiling* vormt. Juist dat is weer aanvechtbaar, maar dat wordt door hen afgedaan met de opmerking: "Hoe dat komt, is een ander verhaal."

Pas na deze twee inleidende paragrafen komen de auteurs met de eerste stap uit het PTOB-schema: het probleem. In deze paragraaf bespreken zij twee ontzuilingsbewegingen. De eerste omschrijven zij (gedeeltelijk tautologisch) als het proces van secularisering én ontzuiling. De tweede als het proces van modernisering (is de secularisering geen vorm van modernisering?). Helaas worden in deze paragraaf geen duidelijke problemen onderscheiden, terwijl deze er natuurlijk wel zijn. Een wetenschappelijk probleem is bijvoorbeeld het vinden van een verklaring voor het feit dat ondanks de sterke secularisering in Nederland, het confessioneel onderwijs vrijwel hetzelfde marktaandeel houdt. Beleidsproblemen zijn de versnippering van het scholenbestand en de extra uitgaven. Een sociaal probleem is, zoals de OESO-examiners opmerkten, dat de vrijheid van onderwijs tot etnische segregatie zou kunnen leiden. De lezer krijgt de indruk dat het vooral om het hierboven genoemde wetenschappelijke probleem gaat: de verklaring voor de achterblijvende *ontzuiling* in het onderwijs. Daarom is het zo merkwaardig dat in de daarop volgende theorie-paragraaf alleen theorieën over *verzuiling* de revue passeren en nergens het verschil tussen *verzuiling* en *ontzuiling* geproblematiseerd wordt.

Tenslotte worden zij in de paragraaf over be-

leid erg normatief. Wat te denken van een zin als: "Weliswaar werkt de tijd in ons voordeel (...), maar weloverwogen beleid verdient de voorkeur boven afwachten." Wie bedoelen de auteurs met 'ons' en waaruit moet zo'n 'weloverwogen' beleid bestaan?

Maken we een balans op van het gehele boek, dan moeten we vaststellen dat deze inleiding veel beter is dan de vorige. De opzet is veel consistent en de gekozen thema's zijn zeker bruikbaar voor studiedoeleinden. Een aantal hoofdstukken is interessant en geeft een helder overzicht van de stand van zaken. Desalniettemin vallen de nodige vraagtekens te plaatsen bij de twee leidraden van het boek. De hoofdvragen worden te weinig beargumenteerd. De onderzoeksproces als leidraad heeft als bezwaar dat sommige thema's in een wat vreemd keurslijf worden geperst en het boek als geheel aan snelle veroudering bloot staat.

S. Karsten

J. J. Hox

Applied Multilevel Analysis

TT-Publikaties, Amsterdam 1995 (second edition)

118 pagina's, f 25,-, ISBN 90-801073-3-6

Dit boek is geschreven ter verheldering van de bedoeling van multiniveau analyse en het werken met multiniveau-programmatuur. Inleidend wordt gesteld dat de multiniveau-optiek vereist is wanneer te onderzoeken objecten onderling genest zijn. Bijvoorbeeld: leerlingen in klassen, of leerlingen in scholen. Er kunnen dan variabelen op het lagere niveau (leerlingen) en variabelen op het hogere niveau (bijv. scholen) onderkend worden. Meer dan twee niveaus zijn ook mogelijk. Op elk niveau, en ook in interacties tussen variabelen van verschillend niveau, kunnen zich processen voordoen die effect hebben op één afhankelijke variabele of op meer afhankelijke variabelen. Deze multiniveau processen en effecten worden statistisch gemodelleerd en geanalyseerd.

Hoofdstuk 2 gaat in op het algemene twee-niveau model en de analyse-strategie die hierbij gebruikt kan worden. Door middel van een voorbeeld wordt tevens aandacht gegeven, modelmatig én wat betreft interpretatie, aan essentiële onderwerpen als multiniveau-interactie, relevantie en noodzaak van centreren, en standaardisering

van coëfficiënten.

Hoofdstuk 3 is gewijd aan het werken met de computerprogramma's die momenteel in het gebruik het belangrijkste zijn: HLM, VARCL en MLn. Zowel overeenkomsten als verschillen tussen de programma's en het werken hiermee komen aan de orde.

In hoofdstuk 4 worden meer specifieke toepassingen behandeld: meta-analyse en multiniveau-analyse met niet-normale (proportionele) data. Uitwerking van de multiniveau aanpak in structurele modellen vindt plaats in hoofdstuk 5. Voorbeelden zijn: multiniveau factoranalyse en multiniveau padanalyse.

Het accent in het boek van Hox ligt op het begrip van en werken met multiniveau analyse in de onderzoekspraktijk. Enerzijds wordt de (beginnende) gebruiker wegwijs gemaakt in de multiniveau-optiek en enkele vaak voorkomende toepassingsaspecten hiervan. Dit mede gerelateerd aan overeenkomsten en verschillen tussen de belangrijkste computerprogramma's. Heel handig is een floppy disc waarop de uitwerking in het boek praktisch te volgen en te doen is. Anderzijds geeft Hox in de laatste hoofdstukken een verder reikend perspectief op multiniveau analyse, waardoor het begrip van de betekenis en toekomstige ontwikkeling hiervan toeneemt. Te

verwachten is dat het belang en gebruik van multiniveau theorievorming en analyse zullen groeien, omdat veel sociaal-wetenschappelijke en ook onderwijskundige problemen methodologisch gezien multiniveau problemen zijn. Het boek is mede om deze reden zeer lezenswaardig.

De uiteenzetting van Hox is in het bijzonder bruikbaar voor onderzoekers die een grondige eerste kijk willen krijgen op de diverse mogelijkheden wat betreft multiniveau-analyse, de programmatuur in dit verband, en het werken hiermee. Ook voor meer ervaren onderzoekers is het boek aan te raden, omdat het een aantal methodologische aspecten verbonden aan multiniveau theorievorming en onderzoek bespreekt en verheldert. Natuurlijk kan het boek niet vervangend zijn voor de bestudering van de aan de specifieke programma's gelieerde handleidingen, of het uitwerken van multiniveau theorie of opdoen van ervaring met de specifieke programmatuur in eigen onderzoeken.

Ten slotte. Naast een paar textuele foutjes in het boek moet de passage 'interviewers are also longer' (p. 25) zijn '(...) also faster'. Variabele Tel wordt op p. 26 anders gecodeerd dan op p. 24 en p. 28; de codering op p. 26 moet zijn: 0 = face-to-face, 1 = telephone.

T. Mooij

Mededelingen

Geïntegreerd informatiesysteem jeugdonderzoek

De Raad van Europa heeft het initiatief genomen een aanzet te geven tot de bouw van een Europees informatiesysteem met betrekking tot jeugdonderzoek. Dit initiatief sluit naadloos aan bij de wens van SWIDOC en PCOJ om de kwaliteit van de Nederlandse Databank Jeugdonderzoek te verhogen door uitbreiding van de bestaande informatie.

Het ministerie van VWS (Directie Jeugdbeleid) heeft het SWIDOC de opdracht gegeven tot het bouwen van bovengenoemd informatie-systeem.

In februari heeft het SWIDOC enquêtes verstuurd naar een aantal onderzoeksinstellingen en onderzoekers, werkzaam op het terrein van jeugdonderzoek. Mocht u deze enquêtes niet hebben ontvangen maar zou u wel graag in de geïntegreerde database willen worden opgenomen, of nadere informatie wensen, dan kunt u contact opnemen met mevr. E. Dijk (coördinator Europese Jeugddatabase), SWIDOC, Herengracht 410-412, 1017 BX Amsterdam, tel.: 020 6225061.

Promoties

Rijksuniversiteit Groningen
Promovendus: Drs. W.G. Sipma
Titel proefschrift: Orthopedagogische thuisbegeleiding met het Portage Programma Nederland
Promotor: Prof.Dr. H. Nakken
Co-promotor: Dr. B.F. van der Meulen
Datum: 18 april 1996

Universiteit Utrecht
Promovendus: Drs. B. A. M. van de Rijt
Titel proefschrift: Voorbereidende rekenvaardigheid bij jonge kinderen
Promotoren: Prof.Dr. L. M. Stevens en Prof.Dr. A. Treffers
Co-promotoren: Dr. J. E. H. van Luit en Dr. A. H. Pennings
Datum: 10 mei 1996

Universiteit Utrecht
Promovendus: Drs. H. A. A. van Eerde

Titel proefschrift: Kwantiwijzer: Diagnostiek in reken-wiskundeonderwijs
Promotoren: Prof.Dr. E. P. J. M. Elbers en Prof.Dr. K. B. Koster
Co-promotor: Dr. R. A. de Jong
Datum: 24 mei 1996

Rijksuniversiteit Leiden
Promovendus: Drs. B. Bijl
Titel proefschrift: Innovaties in de jeugdhulpverlening
Promotor: Prof.Dr. J. D. van der Ploeg
Datum: 20 juni 1996

Rijksuniversiteit Leiden
Promovendus: Drs. M. E. Jacobs
Titel proefschrift: Gezinsachtergronden van schoolsucces bij brugklasleerlingen
Promotoren: Prof.Dr. G. A. Kohnstamm en Prof.Dr. A. J. J. M. Ruijsenaars
Datum: 25 juni 1996

Inhoud andere tijdschriften

Comenius
16e jaargang, nr. 1, 1996

Thema: Straffen of opvoeden

Straffen of opvoeden? door C. Leonards en C. van Nijnatten
From punishment to education? Changing penal responses to juvenile delinquency in late nineteenth and early twentieth centuries - England and Germany compared, door D. Oberwittler
De zware jongens van Gent: praktijk en betoog omtrent de 'incorrigibles' aan het einde van de negentiende eeuw (1875-1921), door J. Christiaens
Andere tijden, andere opvattingen, ander jeugdstrafrecht, door P. van der Laan en verder:
Over sekseverschillen in de beoordeling van psychische klachten, door F. de Boer
Academische vorming op weg naar een nieuw tijdperk: kanttekeningen bij Hoger Onderwijs in Fasen
Philip J. Idenburg (1901-1995): de laatste generalist, door A. Wesselingh

Tijdschrift voor Onderwijsresearch
21e jaargang, nr. 2, 1996

Onderwijsexpansie of meritocratisering? Een verklaring voor de afgenomen samenhang tussen het ouderlijk sociaal milieu en het hoogst bereikte opleidingsniveau van kinderen, door M. H. J. Wolbers en P. M. de Graaf
 De factorstructuur van schrijfvaardigheid: een toepassing van multilevel analyse, door H. Kuhlemeier en H. van den Bergh
 Een onderzoek naar leerstijlen, door H. C. Schouwenburg
 Gedrag van aanstaande docenten in aandachtende situaties in de klas, door W. Admiraal, Th. Wubbels en F. Korthagen
 De constructie en evaluatie van een videotoets voor de beoordeling van bekwaamheid in professionele gespreksvoering, door G. N. Smit en H. T. van der Molen

Tijdschrift voor Orthopedagogiek
 35e jaargang, nr. 4, 1996

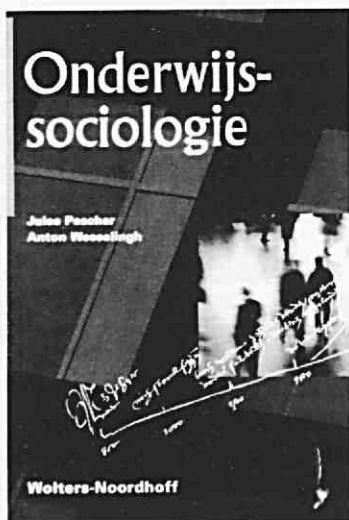
Integratie gehandicapte leerlingen in onderwijs: analyse van vragen voor onderzoek, door M. Søder en S. J. Pijl
 Uitgangspunten en effecten van Videohometraining, door J. M. A. M. Janssens en A. A. M. Kemper
 Kort instrument onderkenning autisme. Verkorting AUTI-R schaal voor gebruik in jeugdgezondheidszorg, door J. Hoekman en J. C. M. M. Duijnhoven

Ontvangen boeken

Meijer, W. A. J., *Stromingen in de pedagogiek*. Uitgeverij Intro, Baarn, 1996, f 29,90.
 Meijvogel, R., *Buitenschoolse opvang. Theorie en praktijk*. Bohn Stafleu Van Loghum, Houten, 1995, f 45,-.
 Neubauer, G., Hurrelmann, K. (Eds.), *Individualization in childhood and adolescence*. W. De Gruyter, Berlin/New York, 1995, DM 148,-.
 Sipma, M., *Orthopedagogische thuisbegeleiding met het Portage Programma Nederland* (dissertatie). Stichting Kinderstudies, Groningen, 1996, f 25,-.
 Weijenberg, J., *Het produktieve groepsteam. Een ontwikkelingsmodel voor groepsteams in de jeugdhulpverlening*. Uitgeverij Van Gorcum, Assen, 1996, f 29,50.

Onderwijssociologie

J. Peschar, A. Wesselingh



Binnen de sociologie heeft zich in Nederland een interessant specialisme ontwikkeld, dat inmiddels grote belangstelling heeft: de onderwijs-sociologie.

Onderwijs is voor de samenleving van zo grote betekenis, dat analyse van de relatie tussen onderwijs en maatschappij noodzakelijk is. Op deze manier kunnen soms verrassende inzichten naar voren komen, die vaak onopgemerkt blijven in de vanzelfsprekendheid van de dagelijkse praktijk.

Het boek Onderwijssociologie is een boeiende introductie op het brede terrein van de onderwijssociologie. Aan de hand van afwisselende onderwerpen wordt de onderwijssociologie overzichtelijk en helder in kaart gebracht.



Bestelcoupon Onderwijssociologie

Ja, ik bestel ----- ex. Onderwijssociologie
ISBN 90 01 706916 à f 59,50

- ☐ rechtstreeks via de uitgever
☐ via boekhandel te -----

☐ de heer ☐ mevrouw

Naam -----

Adres -----

Postcode/Plaats -----

Datum -----

Handtekening -----

Deze coupon kunt u in een ongefrankeerde envelop zenden aan:

Wolters-Noordhoff
t.a.v. de heer J.A. van der Laan
Antwoordnummer 13
9700 VB GRONINGEN
Telefoon (050) 522 65 22

Wolters-Noordhoff



HB: 15 dec 1997
lop

Inhoudsopgave

Pedagogische Studiën Jaargang 73, nummer 4/1996

L. Verhoeven en P. J. J. Stijnen

Aspecten van tweetaligheid: Introductie op het thema 250

P. P. M. Leseman en S. Şahin

Tweetalige ontwikkeling van Turkse peuters 259

R. Aarts, J. J. de Ruiter en L. Verhoeven

Eigen-taalvaardigheid en schoolsucces van Turkse en Marok-
kanse leerlingen 275

S. de Jong en A. M. J. Riemersma

Taalpeiling in Friesland 291

E. B. Klatter

Studievaardigheden in de brugklas. Een vergelijking tussen alge-
mene en vakspecifieke verwerkingsstijlen van brugklasleerlin-
gen 303

Boekbespreking 317

Mededelingen 320

L. Verhoeven en P. J. J. Stijnen

Samenvatting

Voor veel kinderen in Nederland is de taal van het primaire socialisatieproces binnen het gezin een andere dan het Nederlands. Een belangrijke vraag is hoe het (basis)onderwijs het beste kan inspelen op de taalontwikkeling van deze kinderen, zodat hun kansen op schoolsucces en daarmee op maatschappelijk succes worden geoptimaliseerd en tegelijkertijd hun mogelijkheden tot eigen culturele ontwikkeling binnen de Nederlandse samenleving worden vergroot.

In deze bijdrage wordt de ontwikkeling van tweetaligheid nader verkend in een tweetal uiteenlopende verwervingscontexten: Randstad en Friesland. Na een beknopte inleidende oriëntatie op het fenomeen tweetaligheid wordt eerst ingegaan op het verloop van processen van eerste- en tweede-taalverwerving. Vervolgens wordt stilgestaan bij de factoren die de individuele variatie in tweetalige ontwikkeling bepalen. Daarna volgt een reflectie op het beleid inzake het onderwijs aan tweetalige kinderen. Ten slotte worden de artikelen in dit themanummer kort ingeleid.

1 Inleiding

Nederland kent vanouds een meertalige traditie (zie Extra & Verhoeven, 1993a). Behalve het Nederlands worden zowel autochtone als allochtone groepstalen gehanteerd. Naast vele autochtone dialecten geldt het Fries als regionale autochtone groepstaal die door de meerderheid van de bevolking in Friesland als omgangstaal wordt gehanteerd. Daarnaast is er een breed scala aan allochtone groepstalen die van buitenaf in de Nederlandse samenleving zijn geïntroduceerd. Het Chinees en het Italiaans vormen voorbeelden van talen die reeds voor de tweede wereldoorlog werden gebruikt. Het multilinguale karakter van de Nederlandse samenleving is in de naoorlogse jaren in belangrijke mate versterkt als gevolg van de om-

vangrijke immigratie van personen uit (ex)koloniën (Indië, Suriname) en van arbeiders, hun gezinsleden uit met name de Mediterrane landen en van politieke vluchtelingen uit diverse delen van de wereld.

Ondanks de reeds lang bestaande talige heterogeniteit van de Nederlandse samenleving kent het basisonderwijs in Nederland een overwegend monolinguale traditie. Als exclusieve voertaal en doeltaal werd doorgaans het (standaard-)Nederlands gehanteerd. Voor veel kinderen gold (en geldt) dat die taal in sterke of minder sterke mate afweek van de taal van hun primaire socialisatieproces en moest het Nederlands als tweede of als vreemde taal worden geleerd. Mede onder invloed van gebrekkige leerresultaten bij kinderen die van huis uit een andere taal dan het Nederlands spreken werd in de jaren zeventig gezocht naar onderwijsvoorzieningen die beter bij de linguïstische en sociaal-culturele beginsituatie van deze kinderen zouden aansluiten. In Nederland werd aanvankelijk nogal wat aandacht besteed aan de taalsituatie van autochtone kinderen (Friesland (Wijnstra, 1976) en Kerkrade (Stijnen & Vallen, 1981)) (een belangrijk aantal van hun conclusies bleek – later – in versterkte mate van toepassing op allochtone kinderen). Mede als gevolg van de snelle toename in aantal van deze leerlingen werd de taalsituatie van allochtonen in de discussie dominant. Was het onderwijsbeleid aanvankelijk gericht op assimilatie in de Nederlandse samenleving (vgl. het ‘aanspreekbaar’ maken in het Nederlands als doelstelling), geleidelijk groeide het besef dat onderwijs in eigen taal en cultuur eveneens aandacht verdiende.

In Engelstalige immigratielanden als de Verenigde Staten (Veltman, 1983), Canada (De Vries & Vallee, 1980) en Australië (Clyne, 1982) worden periodiek censusdata over thuistaalgebruik en taalvaardigheid verzameld en geanalyseerd. Dergelijke gegevens bieden een rijke bron van informatie over intergenerationale processen van taalverschuiving en taalver-

lies bij uiteenlopende etnische minderheidsgroepen. Voorts kunnen deze gegevens worden gebruikt voor de onderbouwing van overheidsbeleid op vele gebieden van informatieoverdracht (vgl. Extra & Verhoeven, 1993b). In Nederland ontbreekt een dergelijke onderzoekstraditie. Inzicht in de aard en mate van tweetaligheid van de jongste generatie etnische minderheidsgroepen kan belangrijke bouwstenen leveren voor een taalbeleid van de overheid op langere termijn (vgl. Clyne, 1982, pp. 117-147).

Onderzoek naar de ontwikkeling van tweetaligheid van kinderen kan om een aantal redenen worden gemotiveerd. Allereerst is van groot belang zicht te krijgen op de aard en mate van transfer van vaardigheden. Het leren van een tweede taal kan worden gezien als een proces van reorganisatie waarbij eerder verworven declaratieve en procedurele taalkennis als uitgangspunt kan worden genomen voor het opslaan van kennis omtrent de tweede taal. Voorts is van belang inzicht te verkrijgen in de variabelen die de individuele variatie in tweetalige ontwikkeling bepalen. In sociolinguïstisch opzicht lijken vooral de mate van taalcontact en de sociaal-culturele oriëntatie van kinderen variabelen die van belang zijn. Belangrijke onderwijskundige variabelen lijken vooral de mate van instructie in de eerste en de tweede taal en de vormgeving van het curriculum (zie ook Verhoeven en Narain, 1996).

2 Ontwikkeling van tweetaligheid

Met de term 'tweetaligheid' wordt een breed spectrum van taalvaardigheden in twee of meer talen aangeduid. Aanvankelijk sloeg het begrip op taalgebruikers die twee talen op moedertaalniveau beheersen (Bloomfield, 1933). Deze zuivere vorm van tweetaligheid komt echter zelden voor en Bloomfields definitie ('a native-like control of two languages') sloot meer dan de helft van de wereldbevolking uit waar binnen dezelfde taalgemeenschap meerdere talen worden gebruikt (Fishman, 1965). Tweetaligheid wordt tegenwoordig veel ruimer opgevat en omvat het vermogen om twee of meer talen te gebruiken (Mackey, 1968; Romaine, 1989).

Het beheersingsniveau van de twee talen is

niet altijd gelijk en is afhankelijk van onder meer de context, het aanbod en het gebruiksdoel van de taal. Doorgaans worden hierbij drie typen van tweetaligen onderscheiden: tweetaligen die een hoog niveau bereiken in beide talen, tweetaligen bij wie één van de talen dominant is, en tweetaligen bij wie het taalvaardigheidsniveau in beide talen op een laag niveau is blijven steken. In het laatste geval wordt ook wel de term 'semilingualisme' gebruikt, maar de term is omstreden omdat hij stigmatiserend werkt.

Omdat de ontwikkeling van taalvaardigheden domeinafhankelijk is, komt zuivere tweetaligheid nauwelijks voor. Wel wordt de term 'gebalanceerde tweetaligheid' gebruikt voor tweetaligen die beide talen op vrijwel gelijk niveau beheersen. Volgens deze omschrijving worden ook tweetaligen met een laag taalvaardigheidsniveau in beide talen gerekend tot gebalanceerde tweetaligen. Maar volgens vele onderzoekers zijn zelfs gebalanceerde tweetaligen vaardiger in één van de talen, afhankelijk van het domein of de situatie van taalgebruik (Saunders, 1988).

Het is daarom relevant om bij de bestudering van taalvaardigheid van tweetaligen, de aard en mate van tweetaligheid te bepalen. Daarbij moet een onderscheid worden gemaakt tussen een additieve en een subtractieve verwervingscontext (Lambert, 1978). In een additieve context worden een of meer prestige talen toegevoegd aan de eerste taal, die meestal ook de dominante taal is. Vaak is hierbij sprake van een vrije keuze en wordt de term 'elitaire tweetaligheid' gebruikt (Skutnabb-Kangas, 1984). Anders is het in een subtractieve verwervingscontext waarbij de tweede taal de meerderheidstaal is en de eerste taal weinig prestige geniet. Kinderen uit etnische minderheidsgroepen worden meestal in een subtractieve context tweetalig en staan enerzijds onder druk van de omgeving om de meerderheidstaal te leren en anderzijds van het gezin om de eerste taal te handhaven (zie verder voor subtractieve tweetaligheid Leseman & Şahin, dit nummer).

Naast de verwervingscontext, is ook de periode van tweetalige vorming van invloed op de mate van tweetaligheid. Aangenomen wordt dat bij verwerving van een tweede taal in de eerste drie levensjaren van het kind, uitgegaan kan worden van simultane tweetaligheid

(McLaughlin, 1994). Er is sprake van succes-sieve tweetaligheid wanneer de tweede taal wordt verworven nadat de eerste taal reeds tot ontwikkeling is gekomen. Bij kinderen uit etni-sche minderheidsgroepen in Nederland is in de meeste gevallen sprake van successieve twee-taligheid en het T2-aanbod vindt daarbij meest-al plaats op school.

De eerste of eigen taal (T1) is de taal die ge-sproken wordt in de directe sociale omgeving van het kind (meestal het gezin) en wordt vaak de moedertaal of thuistaal genoemd. De eerste-taalontwikkeling van een kind wordt meestal beschreven in opeenvolgende fasen, die zich van elkaar onderscheiden door specifieke ken-merken van taalgebruik (Slobin, 1987). De ont-wikkeling loopt van een prelinguale of voortal-ige fase naar een vroeglinguale fase, een differentiatiefase en tenslotte een voltooiings-fase. In de prelinguale fase in het eerste levens-jaar experimenteert het kind met klanken. In de vroeglinguale periode tussen één en tweeën-half jaar worden woorden gebruikt die bete-kenis hebben, aanvankelijk in één- daarna in tweewoordzinnen. Pas aan het eind van deze fase gaat het kind zinnen gebruiken die uit meer dan twee woorden bestaan. Tot ongeveer het vijfde jaar loopt de differentiatieperiode. In deze fase worden alle verworven vaardigheden verder uitgebouwd, waaronder de fonologie, de woordenschat, de zinsbouw en de morfologie. In deze fase groeien ook de metalinguïst-ische en communicatieve vaardigheden van het kind. Tussen het vijfde en zesde jaar begint het kind taal te gebruiken die de taal van volwasse-nen benadert. Op school leert het kind lezen en schrijven.

Tussen het vermogen om een eerste of een tweede taal te leren is er geen wezenlijk ver-schil, wel in de leercontext. Een kind dat de moedertaal leert, krijgt die taal ongevraagd en frequent aangeboden door een vaste kring van familieleden die steeds dezelfde klanken en woorden herhalen en verbeteren. Hierdoor raakt het kind langzamerhand vertrouwd met de klanken, woorden en structuur van de taal. Een tweede-taalverwerver daarentegen moet vaak eerst sociale contacten (bijvoorbeeld met speelkameraden) opbouwen om de tweede taal te leren. Het aanbod van die tweede taal komt in de basisschoolleeftijd voornamelijk van leef-tijdgenoten en van de leerkracht op school,

maar in geringere mate dan het aanbod van de eerste taal thuis.

In het taalverwervingsproces wordt de aan-geboden taal door de leerder selectief gebruikt en wordt op reeds aanwezige kennis voortge-bouwd. Dat is bij de meeste T2-leerders hun kennis van de eerste taal. Over het algemeen zijn T2-leerders ouder en dus cognitief en so-ciaal verder ontwikkeld dan T1-leerders en be-heersen zij al diverse functies van taal. Aan het begin van het verwervingsproces wordt ge-bruik van grammatische structuren vaak ach-terwege gelaten. Pas in een later stadium wor-den deze structuren verworven volgens nog niet helemaal bekende processen. Omdat alle T2-leerders, ongeacht taalachtergrond of leef-tijd, de structuur van een tweede taal op een-zelfde wijze lijken te leren, wordt aangenomen dat dit proces verloopt op basis van een aange-boren eigenschap van taalleren (Hakuta, 1986). De grote verscheidenheid tussen T2-leerders heeft dan ook minder te maken met de structuur van de taal dan met het tempo van het leren van een tweede taal (Vermeer, 1986). Persoonlijk-heidsfactoren, sociale factoren alsmede indivi-duele verschillen in verbale vaardigheid liggen daaraan ten grondslag.

Cummins (1984) vond empirische evidentie voor een meerdimensionale opvatting van taal-vaardigheid. Hij vond dat de prestaties van leerlingen op een breed scala van taaltaken konden worden samengenomen in twee clus-ters: taken die verwijzen naar algemene cogni-tieve of schoolse kennis en vaardigheden, zoals het schrijven van een informatieve tekst, en communicatieve vaardigheden die gerelateerd zijn aan taalgebruik in interactie. Met betrek-king tot tweetaligheid en tweetalig onderwijs gaat Cummins ervan uit dat meer oppervlak-kige communicatieve taalvaardigheidsaspec-ten min of meer gescheiden tot ontwikkeling komen in een eerste taal en tweede taal (T2). Cognitief/academische vaardigheidsaspecten zouden daarentegen over de twee talen een ge-meenschappelijke basis hebben (zie verder Le-seman & Şahin, dit nummer). Juist die gemeen-schappelijke basis van cognitief/schoolse taalvaardigheid zou een positieve transfer van T1 naar T2 mogelijk maken, zoals verwoord in Cummins' afhankelijkheidshypothese:

"To the extent that instruction in a certain language is effective in promoting profi-

ciency in that language, transfer of this proficiency will occur, provided there is adequate exposure to that other language (either in school or in the environment) and adequate motivation to learn that language”.

Evidentie voor het optreden van taaltransfer is gevonden in een groot aantal studies (voor een overzicht zie Verhoeven, 1994; zie voor ‘cross-linguale transfer’ Leseman & Şahin, dit nummer).

Tweetaligheid heeft geen negatieve invloed op de ontwikkeling van algemene intellectuele vaardigheden van kinderen. Volgens Hakuta (1986) is er juist een positief effect van tweetaligheid op aan taal gerelateerde vaardigheden. Dat geldt met name voor het metalinguïstisch bewustzijn. Onder metalinguïstisch bewustzijn wordt impliciete kennis over taal begrepen die expliciet wordt gemaakt. Bij dit soort kennis verschuift de aandacht van de communicatieve inhoud van taal naar de grammatische vormgeving. Dit gegeven is vooral van belang omdat bij de aanvang van jaargroep 3 het niveau van metacognitieve ontwikkeling van kinderen een belangrijke voorspeller vormt voor het leren lezen.

3 Determinanten van tweetalige ontwikkeling

Onderzoek naar de taalvaardigheid van allochtone kinderen laat zien dat sprake is van een grote individuele variatie in de ontwikkeling van zowel de etnische groepstaal (Verhoeven & Boeschoten, 1986; Verhoeven, 1987; De Ruiter, 1989; Verhoeven, Extra, Konak, Nairn & Zerrouk, 1990) als het Nederlands als tweede taal (Verhoeven & Vermeer, 1989; Boogaard, Damhuis, De Gloppe & Van den Bergh, 1990). Eerder onderzoek naar mogelijke determinanten van de tweetalige ontwikkeling van kinderen van etnische minderheden heeft zich toegespitst op variabelen in de volgende drie domeinen: het kind, het gezin en de institutionele zorg (kindercentra, kleuteropvang) (vgl. Verhoeven, 1991).

Een belangrijke kindvariabele blijkt de culturele oriëntatie van het kind. Culturele oriëntatie is te definiëren als de attitudes die allochtone kinderen hebben ontwikkeld tegenover de

eigen taal en cultuur en de dominante taal en cultuur. Vermeer (1986) en Boers, Van den Bosch en Verhoeven (1987) vonden dat culturele oriëntatie naast taalcontact belangrijke voorspellers van T2-ontwikkeling vormden. Daarnaast blijkt de leeftijd waarop met het leren van een tweede taal een aanvang wordt gemaakt een factor van betekenis. Oudere leerlingen boeken sneller vooruitgang, terwijl jongere leerlingen uiteindelijk succesvoller blijken (vgl. Larsen-Freeman & Long, 1991). Alleen jonge kinderen zijn in staat een accentvrije T2-vaardigheid op te bouwen. Voorts blijkt motivatie een rol van betekenis te spelen. Wong-Fillmore (1976, 1979) liet zien dat de integratieve motivatie van kinderen de T2-ontwikkeling positief beïnvloedde. Kinderen die zich vooral identificeerden met leeftijdgenoten die de T2 spraken boekten de meeste vooruitgang met T2-leren.

Onderzoek naar de invloed van gezinsvariabelen heeft zich vooral gericht op de effectiviteit van het gezin als leeromgeving. Wells (1985) vond dat de mate waarin ouders erin slaan hun taalgebruik af te stemmen op het niveau van het kind van invloed is op de taalontwikkeling van kinderen. De mate van intersubjectiviteit van aandacht, de uitdrukking van begrijpelijke proposities, het toezicht op het succes van de communicatie en de stimulering van verdere interacties bleken daarbij de belangrijkste factoren. Een andere belangrijke gezinsfactor vormt het taalaanbod. Fantini (1985) liet zien dat het taalcontact binnen het gezin van grote invloed is op de taalontwikkeling van allochtone kinderen. Tosi (1979, 1984) toont aan dat het aanbod van T1 over generaties een steeds belangrijker factor wordt met het oog op verwerving van die taal door immigrantenkinderen. Binnen de eerste generatie immigranten is de taal ingebed in een rijke infrastructuur. In latere generaties wordt de eigen taal meestal alleen nog gebruikt in een beperkter aantal domeinen. Een aantal studies laat het belang zien van de mate van contact van gezinsleden met de T2 voor de ontwikkeling van die taal (Schumann, 1978; Vermeer, 1986; Boers, Van den Bosch & Verhoeven, 1987; Verhoeven, 1987). Een andere belangrijke gezinsvariabele vormt de attitudes van ouders tegenover T1-behoud en T2-leren. Gardner liet in de Canadese context zien dat de attitudes van

Engels-sprekende ouders tegenover het leren van Frans als tweede taal de motivatie van kinderen om Frans te leren beïnvloedt. Lalleman (1986) vond eveneens een significant verband tussen de attitudes van ouders en de T2-vaardigheid van jonge Turkse kinderen in Nederland.

Onderzoek laat zien dat institutionele opvang de taalontwikkeling van jonge kinderen in sterke mate beïnvloedt. In een groot aantal studies komt naar voren dat de aard en mate van interactiepatronen in institutionele omgevingen uitermate belangrijk zijn voor de T2-ontwikkeling (vgl. McLaughlin, 1985, pp. 145-163). Met betrekking tot de institutionele opvang van met name jonge allochtone kinderen wordt vaak een tweetalige opvang bepleit. Een groot aantal studies in Canada wijst erop dat kinderen binnen de context van een tweetalig programma een positief zelfbeeld ontwikkelen (vgl. Harley & Lapkin, 1984; Cummins & Swain, 1986).

4 Tweetaligheid en onderwijsbeleid

In Nederland kan het taalbeleid ten aanzien van de Friese gemeenschap worden gekarakteriseerd als een lange reeks van pogingen om het Fries als autochtone taal te erkennen. In 1980 heeft dit geleid tot de verplichtstelling van de Friese taal op alle basisscholen in Friesland. Was het onderwijs in het Fries aanvankelijk bedoeld om Friestalige leerlingen via de Friese taal goed Nederlands te leren, geleidelijk aan werd de doelstelling verruimd. Op dit moment beoogt het onderwijs in Friesland zowel Friestalige als niet-Friestalige leerlingen het Nederlands en het Fries goed te leren beheersen.

Ten aanzien van allochtone groepen (immigranten) was het onderwijsbeleid in Nederland aanvankelijk exclusief gericht op assimilatie met de Nederlandse samenleving. Er werden extra faciliteiten geboden in de vorm van taakleerkrachten die de anderstalige leerlingen in het Nederlands zodanig 'aanspreekbaar' moesten maken dat zij in staat zouden zijn het Nederlandstalige onderwijs te volgen. Geleidelijk aan ging men inzien dat om recht te doen aan ontwikkeling, c.q. behoud van de eigen ervaringswereld, er ook aandacht besteed diende te

worden aan de eigen taal en cultuur. In het 'Beleidsplan culturele minderheden in het onderwijs' (1981) werd het recht op onderwijs in de eigen taal en cultuur (OETC) voorbehouden aan kinderen uit de Molukken, Griekenland, Italië, Joegoslavië, Kaapverdië, Marokko, Portugal, Spanje, Tunesië en Turkije. Het OETC is inmiddels volgens de wet beperkt tot 2 1/2 uur per week onder schooltijd en 2 1/2 uur buiten schooltijd. De middelen voor OETC waren van meet af aan beperkt. Er was geen eenduidig leerplan en er waren geen adequate leermiddelen en toetsen.

Vervolgens is de aandacht gericht op de taalcomponent binnen het eigen taal- en cultuuronderwijs, hetgeen leidde tot het schrappen van de C van het OETC. Er werd betoogd dat met het taalonderwijs de culturele component vanzelf wel de nodige aandacht zou krijgen. Met het aantreden van staatssecretaris Ginjaar-Maas in 1986 ging in de discussies over het OET meer en meer de eigen taal centraal staan en meer en meer werden linguïsten geraadpleegd (Lucassen & Köbben, 1992). In een beleidsnotitie (zie voor een samenvatting Triescheijn, 1989) werd er een grotere nadruk gelegd op de relatie van de eigen taal en het (succesvol) leren van het Nederlands en het belang van de kennis van de eigen taal. De eigen taal zou volgens deze notitie een gevoel van veiligheid geven (Lucassen & Köbben, 1992; zie voor een gelijksoortige conclusie voor autochtone kinderen Stijnen, 1975 en Kuijper, Stijnen & Van den Hoogen, 1983). Deze lijn werd door haar opvolger, staatssecretaris Wallage, overgenomen. In zijn nota *Onderwijs in de Eigen Taal als onderdeel van een geïntegreerd talenonderwijs* (september 1990) stelt hij dat het eigen-taalonderwijs moet worden ingepast in het taalbeleid van de school en dat het Nederlands als tweede taal moet worden onderwezen. In een interview met de schrijvers van 'Het partiële gelijk' (Lucassen & Köbben, 1992) stelt Wallage dat de eigen identiteit van allochtone leerlingen maar beter sterk ontwikkeld kan worden om later beter te kunnen integreren in de Nederlandse maatschappij. In zijn uiteindelijke nota *Eigen Taal* (februari 1991) pleit hij dan ook voor het eigen taalonderwijs als middel tot een versterking van het zelfbeeld van de allochtone kinderen.

De proportie Turkse en Marokkaanse kin-

deren die in het basisonderwijs deelneemt aan het OET blijkt in de afgelopen 15 jaar duidelijk toe te nemen. Was dit aandeel in het schooljaar 1982/1983 nog 60% (Teunissen, 1986), in de schooljaren daarna benaderde het cijfer de 70% (Driessen, De Bot & Jungbluth, 1989). Met name Marokkaanse leerlingen blijken ook buitenschools eigen taalonderwijs te volgen. Shadid en Van Koningsveld (1990a) rapporteren dat in Rotterdam ca. 40 % van de Marokkaanse kinderen tussen de 6 en de 14 jaar moskeewonderwijs volgt. Driessen (1990) registreerde in zijn onderzoek een deelname aan buitenschools 'OETC' van 44% voor de Marokkaanse kinderen en 20% voor de Turkse leerlingen. Over de inhoud van de buitenschoolse lessen is niet veel bekend. Shadid en Van Koningsveld wijzen er echter op dat men niet bang hoeft te zijn voor vormen van islamitisch fundamentalisme. Veeleer kan gesteld worden dat 'dit type onderwijs er is om kinderen in staat te stellen om te zijner tijd als volwaardig lid van de geloofsgemeenschap van hun ouders te functioneren.' Tenslotte zijn er onlangs in Nederland een aantal islamitische scholen opgericht. In het schooljaar 1991-1992 bestonden er in Nederland 20 islamitische scholen (Shadid & Van Koningsveld, 1990b, 1992a). Op deze scholen zaten 2.612 leerlingen, allen van islamitische afkomst; 32,5% is van Turkse en 41,1% van Marokkaanse afkomst. Dat zijn dan respectievelijk 2,5% en 3,1% van de totale groep Turkse en Marokkaanse leerlingen. De instructietaal in deze scholen is het Nederlands maar binnen het religieuze onderwijs is er meer variatie (Shadid & Van Koningsveld, 1992b).

Lag binnen het overheidsbeleid ten aanzien van allochtone groepen het accent nadrukkelijk op maatregelen ter bestrijding van achterstand en achterstelling, recentelijk is er ook aandacht voor cultuurbeleid dat erop gericht is etnische groepen de mogelijkheid te bieden tot eigen culturele ontwikkeling in relatie tot de ander binnen het raamwerk van de Nederlandse samenleving. Onlangs is door de Commissie Allochtone Leerlingen in het Onderwijs onder leiding van Van Kemenade een advies over het onderwijs aan allochtone leerlingen aan de minister uitgebracht (zie Extra, 1994). In dit rapport wordt een beleid voorgestaan dat voorwaarden schept voor culturele pluriformiteit en gelijke kansen voor iedere inwoner in Neder-

land. De belangrijkste doelstellingen daarbij vormen participatie van allochtonen aan de instituties in onze samenleving en aan de samenleving als geheel, en emancipatie in eigen cultuurbeleving binnen de basiswaarden van onze grondwet (Van Kemenade e.a., 1992, pp. 12-13). Een belangrijke implicatie van deze doelstellingen is dat allochtonen worden erkend in hun culturele verscheidenheid. Daarmee dienen zij gelijkkelijk toegang te krijgen tot scholing en arbeidsmarkt en voldoende mogelijkheden te krijgen voor beïnvloeding van de besluitvorming.

In haar aanbevelingen legt de Commissie grote nadruk op het belang van voorschoolse programma's waarin het leren van het Nederlands, het verwerven van basisvaardigheden en het verwerven van sociale redzaamheid centraal staan. Ten aanzien van lessen OET wordt erop gewezen dat marginalisering van dit onderwijs moet worden tegen gegaan door het dezelfde behandeling te geven als andere vakken in de school, inclusief een beoordeling van behaalde prestaties. Bovendien zou het onderwijs in allochtone levende talen in beginsel ook toegankelijk moeten zijn voor autochtone leerlingen (zie in dit verband ook de in dit tijdschrift gevoerde discussie tussen Extra, 1994 en Driessen, 1994).

5 Tweektaligheid onderzocht: introductie op de artikelen

Het thema 'tweektaligheid en basisonderwijs' dat in dit nummer centraal staat, kan op uiteenlopende wijzen worden belicht. Zo kan de nadruk worden gelegd op politieke en juridische facetten van onderwijswetgeving, op culturele, sociaal-psychologische en taalkundige aspecten van het gebruik van talen binnen het gezin en op school, dan wel op pedagogische en onderwijskundige voorstellen voor de inrichting van het onderwijs aan tweektalige kinderen. Verder is duidelijk dat de problematiek van tweektaligheid in relatie tot onderwijs sterk afhankelijk is van de sociaal-politieke context waarin kinderen leven. Vanuit een multidisciplinair kader wordt het thema in de navolgende artikelen toegespitst op de ontwikkeling van tweektaligheid bij kinderen in de volgende twee gebieden:

- de Randstad waar allochtone minderheden van huis uit overwegend een eigen etnische groepstaal hanteren en in het onderwijs vanuit het Nederlands instructie krijgen met daarnaast enkele uren per week les in de eigen taal;
- Friesland, waar het Fries als autochtone taal wordt gesproken naast het Nederlands en in het onderwijs Nederlands de instructietaal is en het Fries als vak wordt onderwezen.

Daarbij wordt aandacht geschonken aan de periode direct voorafgaande aan het basisonderwijs, de basisonderwijsperiode en het begin van het vervolgonderwijs.

Leseman en Şahin bespreken het proces van tweetalige ontwikkeling van Turkse peuters in Nederland. Vanuit longitudinaal perspectief gaan zij na hoe de woordenschatontwikkeling van 31 Turkse kleuters in twee talen verloopt in vergelijking met die van 77 Nederlandstalige leeftijdgenoten. Daarbij wordt ook evidentie voor het optreden van taaltransfer gezocht. Daarnaast relateren zij het tempo van tweetalige ontwikkeling aan een reeks van achtergrondkenmerken. De resultaten laten een grote mate van dominantie in de eigen taal van de Turkse peuters zien. Over het geheel genomen blijkt de taalontwikkeling van de Turkse kinderen (in het Turks) niet substantieel af te wijken van een vergelijkingsgroep van Nederlandstalige kinderen uit dezelfde sociaal-economische klasse. Verder lijkt de taalontwikkeling in het Turks ten koste te gaan van de (versnelling in de) taalontwikkeling in het Nederlands.

Aarts, De Ruiter en Verhoeven gaan in op de relatie tussen taalvaardigheid en schoolsucces bij Turkse en Marokkaanse leerlingen op het einde van het basisonderwijs. Het onderzoek richt zich met name op de overgang van basis naar voortgezet onderwijs, die van groot belang voor de verdere schoolloopbaan van leerlingen kan worden beschouwd. De rol van de eigen taal van leerlingen krijgt daarbij een centrale plaats toebedeeld. Nagegaan wordt allereerst welke factoren de individuele variatie van leerlingen op Cito-Eindtoetsscores verklaren. Verder wordt bepaald in hoeverre het advies voor vervolgonderwijs van de klasseleerkracht kan worden voorspeld vanuit de beheersing van leerlingen van de eigen taal. Ook ten aan-

zien van de feitelijke doorstroming van leerlingen in het voortgezet onderwijs wordt de rol van de eigen taal onderzocht.

Riemersma en De Jong ten slotte geven in hun artikel een beschrijving van een gehouden taalpeiling in Friesland. Daarin is bij kinderen op het einde van de basisschool de vaardigheid in zowel het Fries als het Nederlands bepaald. De opzet van het onderzoek is gebaseerd op een eerder door het Cito uitgevoerd periodiek peilingsonderzoek. Op deze wijze is het mogelijk om, wat het Nederlands betreft, een vergelijking te maken met een landelijke normgroep. De auteurs komen tot de conclusie dat leerlingen in Friesland op het eind van het basisonderwijs wat hun beheersing van het Nederlands betreft even hoog uitkomen als andere leerlingen in Nederland, een bevinding die ook Wijnstra in 1976 en Stijnen en Vallen voor Kerkraadse dialectsprekers in 1981 rapporteerden. De beheersing van het Fries blijkt volgens Riemersma en De Jong echter niet te voldoen aan in de kerndoelen gestelde verwachtingen.

Literatuur

- Bloomfield, L. (1933). *Language*. New York: Holt.
- Boers, M., Bosch, L., van den, & Verhoeven, L. (1987). Taalaanbod en culturele oriëntatie als voorspelers van tweede-taalvaardigheid. *Tijdschrift voor Taalbeheersing*, 9, 166-180.
- Boogaard, M., Damhuis, R., Glopper, K., de, & Bergh, H., van den (1990). *De Nederlandse taalvaardigheid van allochtone en Nederlandse kleuters*. Lisse: Swets & Zeitlinger.
- Clyne, M. (1982). *Multilingual Australia. Resources, needs, policies*. Melbourne: River Seine Publ.
- Cummins, J. (1984). Wanted: a theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 2-19). Clevedon: Multilingual Matters.
- Cummins J., & Swain, M. (1986). *Bilingualism in education*. London: Longman.
- Driessen, G. (1990). *De onderwijspositie van allochtone leerlingen*. Nijmegen: ITS.
- Driessen, G. (1994). Naar een meer realistische benadering van het Onderwijs in Eigen Taal? *Pedagogische Studiën*, 71, 47-59.

- Driessen, G., Bot, K. de, & Jungbluth, P. (1989). *De effectiviteit van het Onderwijs in Eigen Taal en Cultuur. Prestaties van Marokkaanse, Spaanse en Turkse leerlingen*. Nijmegen: ITS.
- Extra, G. (1994). Etnische minderheden, taaldiversiteit en onderwijs in eigen taal. *Pedagogische Studiën*, 71, 35-46.
- Extra, G., & Verhoeven, L. (1993a). *Community languages in the Netherlands*. Lisse: Swets & Zeitlinger.
- Extra, G., & Verhoeven, L. (1993b). *Immigrant Languages in Europe*. Clevedon: Multilingual Matters.
- Fantini, A. (1985). *Language acquisition of a bilingual child*. Clevedon: Multilingual Matters.
- Fishman, J. (1965). Bilingualism, intelligence and language learning. *The Modern Language Journal*, 4, 227-236.
- Hakuta, R. (1986). *Mirror of language: The debate on bilingualism*. New York: Basic Books.
- Harley, B., & Lapkin, S. (1984). *The effects of early bilingual schooling on first language development*. Toronto: OISE.
- Kemenade, J.A. van (red.). (1992) *Ceders in de tuin*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.
- Kuijper, H., Stijnen, S., & Hoogen, J. van den (1983). *Onderwijs tussen dialect en standaardtaal*. Nijmegen: KUN/NCDN.
- Lalleman, J. (1986). *Dutch language proficiency of Turkish children born in The Netherlands*. Dordrecht: Foris.
- Lambert, W. (1978). Psychological approaches to bilingualism, translation and interpretation. In D. Gerver & H. Sinaiko (Eds.), *Language interpretation and communication* (pp. 131-144). New York: Plenum Press.
- Larsen-Freeman, D., & Long, M. (1991). *An introduction to second language acquisition research*. London: Longman.
- Lucassen, J., & Köbben, A. J. F. (1992). *Het partiële ge-lijjk. Controverses over het onderwijs in de eigen taal en cultuur en de rol daarbij van beleid en wetenschap*. Lisse: Swets & Zeitlinger.
- Mackey, W. (1968). The description of bilingualism. In J. Fishman (Ed.), *Readings in the Sociology of Language*. The Hague: Mouton.
- McLaughlin, B. (1985). *Second language acquisition in childhood*. Hillsdale, NJ: Erlbaum.
- McLaughlin, B. (1994). Retrospective on the study of early bilingualism. In G. Extra & L. Verhoeven (Eds.), *The cross-linguistic study of bilingual development* (pp. 27-38). Amsterdam: North Holland.
- Romaine, S. (1989). *Bilingualism*. Oxford: Basil Blackwell.
- Ruiter, J.J. de (1989). *Young Moroccans in the Netherlands. An integral approach to their language situation and acquisition of Dutch*. Utrecht: Academisch proefschrift RUU.
- Saunders, G. (1988). *Bilingual children: from birth to teens*. Clevedon: Multilingual Matters.
- Schumann, J. (1978). The acculturation model for second language acquisition. In R. Gingras (Ed.), *Second language acquisition and foreign language teaching*. Arlington: Center for Applied Linguistics.
- Shahid, W. A. R., & Koningsveld, P. S. van (1990a). *Muslims in Nederland; minderheden en religie in een multiculturele samenleving*. Alphen aan den Rijn: Samsom Stafleu.
- Shahid, W. A. R., & Koningsveld, P. S. van (1990b). *Islamitische basisscholen in Nederland. Samenwijken*, 11, 19-23.
- Shahid, W. A. R., & Koningsveld, P. S. van (1992a). *Islamic primary schools*. In W. A. R. Shahid & P. S. van Koningsveld (Eds.), *Islam in Dutch society: current developments and future prospects*. Kampen: Pharos.
- Shahid, W. A. R., & Koningsveld, P. S. van (1992b). *De mythe van het Islamitische gevaar: hindernissen bij integratie*. Kampen: Kok.
- Skutnabb-Kangas, T. (1984). *Bilingualism or not. The education of minorities*. Clevedon: Multilingual Matters.
- Slobin, D. I. (1987). *The cross-linguistic study of language development*. Hillsdale, NJ: Erlbaum.
- Stijnen, P. J. J. (1975). *Het van huis uit spreken van dialect en onderwijs. Een literatuurstudie*. Nijmegen: NIVOR.
- Stijnen, S., & Vallen, T. (1981). *Dialect als onderwijsprobleem. Een sociolinguistisch-onderwijskundig onderzoek naar problemen van dialectsprekende kinderen in het basisonderwijs*. Academisch proefschrift. 's-Gravenhage: Staatsuitgeverij.
- Teunissen, F. (1986). *Een school, twee talen*. Academisch proefschrift. Utrecht: RUU.
- Tosi, A. (1979). Mother tongue teaching for the children of migrants. *Language Teaching and Linguistics: Abstracts*, 12, 213-231.

- Tosi, A. (1984). *Immigration and bilingual education*. Oxford: Pergamon Press.
- Triesscheijn, T. J. M. (1989). Staatssecretaris voor meertaligheid. *Samenwijs*, 10, 72-75.
- Veltman, C. (1983). *Language shift in the United States*. Berlin: Mouton.
- Verhoeven, L. (1987). *Ethnic minority children acquiring literacy*. Published Ph. D. dissertation. Berlin: Mouton.
- Verhoeven, L. (1991). Predicting minority children's bilingual proficiency: Child, family and institutional factors. *Language Learning*, 41, 205-233.
- Verhoeven, L. (1994). Transfer in bilingual development. *Language Learning*, 44, 381-415.
- Verhoeven, L., & Boeschoten, H. (1986). First language acquisition in a second language environment. *Journal of Applied Psycholinguistics*, 7, 241-255.
- Verhoeven, L., Extra, G., Konak, O., Narain, G., & Zerrouk, R. (1990). *Vroege tweetaligheid bij Turkse, Marokkaanse en Antilliaanse/Arubaanse kleuters*. Tilburg: Tilburg University Press.
- Verhoeven, L., & Narain, G. (1996). Ontwikkeling van tweetaligheid bij allochtone kleuters. *Tijdschrift voor Onderwijsresearch*, 21, 54-80.
- Verhoeven, L., & Vermeer, A. (1989). *Diagnose van kindertaal*. Tilburg: Zwijzen.
- Vermeer, A. (1986) *Tempo en structuur van tweede-taalverwerving bij Turkse en Marokkaanse kinderen*. Academisch proefschrift. Tilburg: KUB.
- Vries, J. de, & Vallee, F. (1980). *Language use in Canada*. Ottawa: Ministry of Supply and Services.
- Wells, G. (1985). *Language Development in the Pre-school Years*. Cambridge: University Press.
- Wijnstra, J. M. (1976). *Het onderwijs aan van huis uit friestalige kinderen. Verslag van een evaluatie-onderzoek in een meertalige regio*. Academisch proefschrift. 's-Gravenhage: Staatsuitgeverij.
- Wong-Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Published Ph. D. dissertation. Stanford, CA: Stanford University.
- Wong-Fillmore, L. (1979). Individual differences in second language learning. In C. J. Fillmore, W. S. Wang & D. Kempler (Eds.), *Individual differences in language ability and language behavior* (pp. 283-296). New York: Academic Press.

Auteurs

L. Verhoeven is werkzaam als hoogleraar Orthopedagogiek aan de Katholieke Universiteit Nijmegen.

P. J. J. Stijnen is werkzaam als universitair hoofddocent bij de faculteit Sociale Wetenschappen en als Onderwijsdecaan bij de faculteit Cultuurwetenschappen van de Open Universiteit te Heerlen.

Correspondentie-adres: Vakgroep Orthopedagogiek KUN, Montessorilaan 3, 6525 HR Nijmegen

Abstract

Bilingualism and schooling: an introduction

L. Verhoeven & P. J. J. Stijnen. *Pedagogische Studiën*, 1996, 73, 250-258.

The focus of this issue is on bilingualism and schooling. In three articles recent research developments in the field of bilingual development will be discussed in two contexts: the urban part of the Netherlands and Frisia. In this article a short discussion is presented on the theme. First of all, research on first and second language development of children at school age is dealt with. Special attention is given to processes of language transfer that can take place during children's development. In addition, the factors determining the individual variation in first and second language development are examined on the basis of a review of earlier studies. Then, various language policies with respect to the education of children dealing with more than one language are discussed, followed by a short introduction of the three articles.

P. P. M. Leseman en S. Şahin

Samenvatting

In een kleinschalige studie is de ontwikkeling van de actieve en passieve Nederlandse en Turkse woordenschat van 31 Turkse peuters van derde tot vierde verjaardag longitudinaal gevolgd en vergeleken met de Nederlandse taalontwikkeling van 77 even oude Nederlandse peuters. Met behulp van variantie-analyses voor herhaalde waarnemingen zijn hoofdeffekten van 'groep' (Turks, Nederlands-laag SES, Nederlands-hoog-SES) en 'leeftijd' onderzocht. De ontwikkeling van het Turks van Turkse peuters wijkt niet veel af van de ontwikkeling van het Nederlands van Nederlandse laag-SES peuters. De ontwikkeling van het Nederlands van de Turkse peuters blijft echter ver achter. Significante interactie-effecten suggereren dat in de periode vóór en rond de vierde verjaardag, de versnelde ontwikkeling van het Nederlands van Turkse kinderen (door peuterspeelzaalbezoek en intrede in het basisonderwijs) gepaard gaat aan een vertraagde ontwikkeling van het Turks. Een correlatieve analyse met enkele gezinskenmerken en gegevens over het gebruik van Nederlandstalige kinderopvang geeft verdere aanwijzingen over de oorzaken van de vertraagde ontwikkeling van de eigen taal.

1 Inleiding

1.1 De taalsituatie van Turkse kinderen

Turkse kinderen in Nederland groeien meestal op in een meertalige omgeving waarin vier talen of taalvariëteiten kunnen worden onderscheiden (Extra & Vallen, 1990; Van der Heijden, 1992): een dialect van het Turks, standaard Turks, een dialect van het Nederlands en standaard Nederlands. Volgens gege-

vens van Boeschoten, Verhoeven en Vermeer (1986) die gebaseerd zijn op onderzoek in het begin van de jaren tachtig, is voor jonge Turkse kinderen een Turks dialect kwantitatief gezien veruit de belangrijkste variëteit in het taalaanbod. Dat hangt samen met de streek van herkomst van de ouders, meestal een plattelandsg gebied in Turkije, hun scholingsgraad, die meestal laag is, en hun sociale netwerk in Nederland dat vaak bestaat uit familieleden en streekgenoten. De betekenis van standaard Turks in het taalaanbod van jonge Turkse kinderen is meestal veel geringer. Dat hangt enerzijds samen met de al genoemde sociaal-economische en culturele positie van de ouders, anderzijds met het ontbreken van voorzieningen, zoals kinderopvang, speelzalen en, in een latere fase, scholen waar standaard Turks een belangrijke voertaal of instructietaal is.

Nederlands wordt een belangrijk onderdeel van het taalaanbod op het moment dat kinderen naar school gaan. Voor zover bekend gaat het overgrote deel van de Turkse kinderen kort na hun vierde verjaardag naar de basisschool; vanaf vijf jaar vallen zij onder de leerplicht. Voor die tijd bezoeken Turkse kinderen soms een peuterspeelzaal of maken de ouders gebruik van kinderopvang in een crèche, waar (standaard) Nederlands bijna altijd de enige taal is die de leid(st)ers spreken. Recente onderzoeksgegevens wijzen overigens uit dat Turkse ouders relatief weinig gebruik maken van kinderopvangvoorzieningen (Van den Berg & Vlug, 1993; Leseman, Sijsling, Jap-A-Joe & Şahin, 1995).

Een andere manier waarop het Nederlands in het taalaanbod van jonge Turkse kinderen geleidelijk aan belangrijker wordt, is via oudere schoolgaande broertjes en zusjes en Nederlandstalige vriend(innet)jes in de buurt. Behalve dat meestal sprake zal zijn van een dialect van het Nederlands, bijvoorbeeld Amsterdams, moet bedacht worden dat het taal betreft van kinderen die ten dele zelf nog in het proces van (tweede) taalverwerving verkeren.

* Het onderzoek waarover hier gerapporteerd wordt is gesubsidieerd door de Stichting Sociaal-Culturele Wetenschappen (thans SCR) van NWO.

1.2 Kwaliteit van het taalgebruik

Het onderscheid in dialecten en standaardtalen, en een formele, kwantitatieve inventarisatie van het taalgebruik naar taalvariëteiten, is één manier om de aard van het taalaanbod te beschouwen. Een andere manier is de inhoudelijke kwaliteit van het taalgebruik, in het bijzonder in ouder-kind interacties, te onderzoeken. In dit verband is een ander onderscheid van belang, dat in theorie onafhankelijk is van taalvariëteiten. Enerzijds kan taal gebruikt worden voor cognitief-complexe, informatieve en educatieve functies. Het gaat om gebruik van taal om te communiceren over niet-onmiddellijk aanwezige, soms abstracte en puur hypothetische situaties, oftewel 'gedecontextualiseerd taalgebruik'. Anderzijds wordt taal gebruikt voor instrumentele en informeel-sociale doelen die vaker de onmiddellijke situatie of een grotendeels door de participanten gedeelde wereld betreffen. Deze vorm van taalgebruik kan als relatief 'gecontextualiseerd' worden beschouwd (Cummins, 1984; Wells, 1987; Snow, 1991). Cummins (1984) maakt in dit verband onderscheid tussen *Basic Interpersonal Communicative Skills*, verwijzend naar de kennis en vaardigheden die betrokken zijn bij gecontextualiseerd taalgebruik, en *Cognitive Academic Language Proficiency*, de vaardigheid om taal op een gedecontextualiseerde manier te gebruiken.

De verhouding tussen deze twee categorieën taalgebruik in het taalaanbod van een kind is van fundamentele invloed op de taalontwikkeling (Olson, Bayles & Bates, 1986; McDevitt, Hess & Kashiwagi, 1987). Verschillende onderzoeken tonen dat de verhouding van contextgebonden en gedecontextualiseerd taalgebruik in het taalaanbod van een gezin (en het bredere familie-/sociale netwerk) sterk samenhangt met de sociaal-economische positie van het gezin, dat wil zeggen: met de scholingsgraad en het beroepsniveau, en met culturele kenmerken als de geletterdheid van de volwassenen, afgemeten aan de tijd die zij aan het lezen van boeken en kranten besteden (Heath, 1983; DeTemple & Beals, 1991; Leseman, 1994).

Over de hier bedoelde kwalitatieve dimensie van het taalaanbod van jonge Turkse kinderen zijn weinig directe gegevens voor handen (zie echter Aarts & Van Baren, 1990; Huls

& Van de Mond, 1992; Leseman, et al., 1995). Desalniettemin kan, gelet op de sociaal-economische positie van Turkse gezinnen in Nederland en de nog altijd relatief hoge graad van functionele ongeletterdheid (cf. Doets, 1992), vermoed worden dat het taalaanbod *inhoudelijk* sterk door contextgebonden taalgebruik wordt gekenmerkt.

1.3 Kernvragen in het tweetaligheids-onderzoek

Onderzoek en theorievorming op het gebied van tweetalige ontwikkeling geeft geen eenduidig beeld van de gevolgen van een tweetalig taalaanbod op de ontwikkeling van de beide talen afzonderlijk (zie bijvoorbeeld Cummins, 1981, 1984; Hakuta, 1986; Durán, 1987; Ogbu, 1992; Kook, 1994). Twee begrippen die in dit verband van belang zijn en tot op zekere hoogte tegenover elkaar gezet kunnen worden, zijn 'subtractieve tweetaligheid' en 'cross-linguale transfer'. Met subtractieve tweetaligheid wordt bedoeld op het fenomeen dat de verwerving en het gebruik van de ene taal (meestal de tweede, die de meerderheidstaal is) *ten koste* gaat van de ontwikkeling en het gebruik van de andere taal (meestal de moedertaal van taalminderheidsgroepen). Als verklaring hiervoor wordt primair op sociolinguïstische en sociaal-psychologische oorzaken gewezen. De ongelijke maatschappelijke status van meerderheids- en minderheidstaal gaat gepaard met een neiging om de eigen minderheidstaal op te geven of in elk geval uitsluitend te gebruiken in informele situaties; omgekeerd zou een sterke identificatie met de eigen taal en cultuur verwerving van de meerderheidstaal belemmeren. Als in een samenleving geen formeel aanbod van de minderheidstaal wordt gecreëerd voor kinderen, bijvoorbeeld in het onderwijs, dreigt de ontwikkeling van de eigen taal te stagneren.

Het andere begrip, cross-linguale transfer, verwijst naar het feit dat ontwikkeling van de vaardigheid in de ene taal *voordeel* biedt aan de verwerving van de andere taal. Hiervoor dienen echter wel geschikte omstandigheden gecreëerd te worden, bijvoorbeeld door kinderen de kans te geven zich in de eigen minderheidstaal te ontwikkelen voordat met onderwijs in de meerderheidstaal wordt begonnen, of door meerderheids- en minderheidstaal in het onderwijs als gelijkwaardig te behandelen en ruw-

weg in gelijke mate en op gelijk niveau aan te bieden (Verhoeven, 1993).

Onder welke condities subtractie en transfer optreden in meertalige situaties is, naast een sociale en cultuur-politieke, ook een cognitief-(taal)psychologische vraag. Cognitief-psychologische verklaringen voor subtractie-verschijnselen of pendanten daarvan (interferentie, codeverwisseling, taalvereenvoudiging, lage beheersingsniveaus in beide talen, e.d.) worden gezocht in de wijzen waarop taalspecifieke, algemene taal- en taal-onafhankelijke kennis en vaardigheden mentaal worden gerepresenteerd. Een gangbare opvatting is dat de mentale opslag van twee talen grotendeels gescheiden is, maar op het niveau van algemene (taal)kennis en vaardigheden gedeeld wordt (cf. Levelt, 1989).

Transfer verwijst in dit globale cognitieve model naar het leggen van kruisverbanden tussen beide taalspecifieke cognitieve structuren via de taalafhankelijke kennis- en vaardigheidsrepresentatie. Transfer tussen twee talen is aangetoond zowel op het niveau van semantische woord- en begripkennis, als op het niveau van taalanalytische en metalinguïstische vaardigheden (zie onder meer Verhoeven & Van Kuyk, 1991; Narrain & Verhoeven, 1993; Kook, 1994). Interferentie-verschijnselen zouden volgens het cognitieve model onder andere kunnen wijzen op niet-geslaagde specifieke taalopslag, waarvoor de oorzaken vooral in de aard van het taalaanbod gezocht moeten worden. In dit verband merkt Cummins (1981) op dat veelvuldige wisselingen van taalcode in het taalaanbod zonder dat dit duidelijk gemarkeerd is door het onderwerp van gesprek, door de personen waarmee of de situaties waarin de betreffende taal wordt gesproken, negatieve gevolgen heeft voor de taalontwikkeling. Sommige auteurs bevelen om die reden aan vooral bij jonge kinderen vast te houden aan de zogenaamde 'één taal, één persoon'-strategie.

Tweetalige taalverwerving en transfer van kennis van een taal naar een andere zijn waarschijnlijk ten dele 'interne', spontaan verlopende cognitieve processen, die zo bezien alleen maar vergen dat het tweetalige kind voldoende contact heeft met beide talen. De algemene intelligentie van het kind zal hierbij een belangrijke bron van individuele variatie zijn. Veel aanknopingspunten voor een effec-

tieve didactiek biedt deze analyse echter niet. Ook is het de vraag of hiermee verschillen in de ontwikkeling van tweetaligheid en cross-linguale transfer die samenhangen met de sociaal-economische positie en andere gezinskenmerken, wel afdoende verklaard kunnen worden.

Kook (1994) bespreekt de verschillende benaderingen in tweetaligheidsonderzoek tegen de achtergrond van cognitief-psychologische theorievorming over transfer en Vygotskij's theorie over de sociaal gedetermineerde ontwikkeling van hogere cognitieve vaardigheden. Zij veronderstelt dat bevordering van transfer in sociale interacties, doordat een ouder of leerkracht op overeenkomsten wijst, analogieën aandraagt, woorden vertaalt en metalinguïstische informatie geeft, essentieel is voor de ontwikkeling van evenwichtige, niet-subtractieve tweetaligheid. Om sociaal-cultureel (of onderwijskundig) bepaalde verschillen in tweetalige ontwikkeling te kunnen verklaren, biedt deze veronderstelling een goed aanknopingspunt. Wat betreft de effecten van tweetaligheid, wordt de aandacht mede gericht op (tweetalige) taalinteractie en taal instructie.

1.4 Tweetalige ontwikkeling van Turkse kinderen: eerder onderzoek

Recente Nederlandse onderzoeken, met name onder Turkse schoolkinderen, en vergelijkende studies onder Turkse kinderen in Nederland en Turkije lijken uit te wijzen dat de taalsituatie voor Turkse kinderen in Nederland, kort samengevat in de inleiding van dit artikel, ongunstig is voor zowel hun Nederlandse als Turkse taalontwikkeling (zie bijvoorbeeld Appel & Schaufeli, 1990; Narrain & Verhoeven, 1993). Er is met andere woorden sprake van een situatie van 'subtractieve tweetaligheid'.

We vatten de bevindingen kort samen. Turkse kinderen in Nederland komen het onderwijs binnen met een meestal veel geringere kennis van het Nederlands dan hun Nederlandse leeftijdsgenootjes (zie ook Boogaard, Damhuis, Van den Bergh & De Gloppe, 1991; De Jong, Klapwijk & Van der Leij, 1995). De kennis van hun eigen taal is op dat moment veel verder ontwikkeld dan hun kennis van het Nederlands, en doet weinig of niets onder voor de kennis die (qua sociaal-economische positie) vergelijkbare Nederlandse kinderen hebben

van het Nederlands, hún eigen taal.

Na verloop van tijd verandert het patroon: de Nederlandse taalvaardigheid gaat veel sneller vooruit dan de Turkse taalvaardigheid, en heeft weldra (tegen het einde van de basisschool) een hoger niveau bereikt: de kinderen zijn op dat moment dominant in de tweede taal. De achterstand in Nederlandse taalvaardigheid ten opzichte van Nederlandse kinderen wordt in deze periode weliswaar veel kleiner, maar het aanvankelijke verschil wordt gedurende de basisschool slechts gedeeltelijk ingelopen.

Verhoeven en Boeschoten (1986) vergeleken de taalontwikkeling in het Turks van Turkse kinderen in Nederland met die van naar sociaal-economische positie vergelijkbare Turkse kinderen in Turkije. Op vijfjarige leeftijd was er nog geen verschil in ontwikkelingsniveau, maar een paar jaar later was er al een flinke achterstand van de Nederlands-Turkse kinderen. Het lijkt weinig twijfel dat de dominantie van de Nederlandse taal in het onderwijs hieraan debet is, maar ook valt af te leiden dat het informele aanbod van Turks in de buitenschoolse context, met name het gezin, hiervoor kennelijk niet kan compenseren.

2 Doel van het onderzoek

Over de tweetalige ontwikkeling van Turkse kinderen die al op de basisschool zitten is inmiddels het nodige bekend. Gegevens over het aanbod Nederlands waarmee Turkse peuters in contact komen voordat zij naar de basisschool gaan, bijvoorbeeld via Nederlands taalaanbod in het gezin, de peuterspeelzaal of het dagverblijf dat zij bezoeken, zijn echter schaars of gestoeld op kwalitatieve gevalsstudies. Evenmin is er veel inzicht in de effecten van de tweetalige situatie op de ontwikkeling van het Turks en het Nederlands van Turkse peuters in de voorschoolse periode, noch in de (moderere) effecten van sociaal-economische en pedagogische determinanten van (tweetalige) taalontwikkeling.

Het longitudinale onderzoek waarover hierna gerapporteerd wordt, beoogt het inzicht in de voorschoolse taalontwikkeling van Turkse peuters te vergroten en tracht in het bijzonder na te gaan of de taalsituatie waarin Turkse peuters opgroeien, bevorderend dan

wel belemmerend is voor evenwichtige tweetalige ontwikkeling. Het onderzoek is overigens niet als taalverwervingsonderzoek opgezet, maar vertrokken vanuit een algemenere onderwijssociologische vraag naar vroege sociaal-culturele oorzaken van onderwijsongelijkheid (zie Leseman, Vergeer, Sijsling, Jap-A-Joe & Şahin, 1992; Leseman, 1993). Taalontwikkeling, tweetalige ontwikkeling van de Turkse kinderen in het bijzonder, vormt hierbij uiteraard een belangrijk aandachtspunt.

3 Methode

3.1 Onderzoeksofzet en onderzoeksgroep

Het onderzoek is opgezet als een longitudinaal onderzoek met drie meetmomenten met een periode van ongeveer zes maanden tussen de meetmomenten. Op het eerste meetmoment waren de kinderen gemiddeld drie jaar, met een range van vier maanden rond de derde verjaardag. Het laatste meetmoment viel voor de meeste peuters (kort) na de vierde verjaardag, op een moment dat de meesten reeds enkele weken naar de basisschool gingen.

Een deel van de Turkse gezinnen (12) is geworven door middel van een mailing op basis van een adressenbestand van de gemeente Utrecht. De adressen (in totaal 67) waren getrokken uit het bevolkingsregister, geselecteerd naar nationaliteit van de ouders en geboortedatum van het doelkind. Vanwege privacybescherming was het niet mogelijk direct in contact te treden met de beoogde respondenten of te rappeleren. De overige gezinnen (19) zijn geworven door bezoek aan huis op basis van een adressenbestand van de gemeente Rotterdam. Er werden 45 adressen bezocht en aanvankelijk 20 gezinnen geworven, wat voor het totale onderzoek voldoende was. Nog voordat het onderzoek was begonnen viel een gezin af (reden onbekend). Op het eerste meetmoment deden 31 Turkse gezinnen mee, op het tweede meetmoment, een half jaar later, konden er nog 28 gezinnen bereikt worden en op het derde meetmoment, weer een half jaar later, 27. De belangrijkste oorzaak voor de uitval was verhuizing naar Turkije (twee gezinnen) en verhuizing naar een ander, niet te achterhalen adres (een gezin). In een geval gaf de vader op het tweede (en derde) meetmoment moeder en

kind geen toestemming om mee te doen.

De vergelijkingsgroep die ten behoeve van de analyses in dit artikel uit een groter cohort is samengesteld bestaat uit Nederlandse peuters van 3 jaar die Nederlands als moedertaal spreken ($n=77$). Overigens bleek in enkele gezinnen door familieleden en kennissen af en toe een andere taal gesproken te worden. De gezinnen zijn allen schriftelijk geworven op basis van een adressenbestand van de gemeente Utrecht. Het positieve responspercentage zonder rappel bedroeg 38%. De selectie naar moedertaal is gebaseerd op gegevens over het taalgebruik thuis en over de moedertaal van de moeder van het kind. De vergelijkingsgroep is verdeeld in een laag-SES ($n=31$) en hoog-SES groep ($n=46$). Van de laag-SES groep is het opleidingsniveau van beide ouders gemiddeld niet hoger dan LBO.

3.2 Metingen

Testafname geschiedde thuis in het bijzijn van de moeder of de vader. De Turkse peuters werden door een Turkse onderzoekster getest. Omdat er behalve de taaltoetsen nog meer toetsen werden afgenomen en een interview plaatsvond met de ouders, werd de testafname over verschillende dagen verspreid. De kinderen konden meestal niet langer dan een half uur achtereen worden getest. Bij alle subtests is de door de testconstructeurs aanbevolen afbreekregel toegepast, mede om frustratie te vermijden en de testtijd zo kort mogelijk te houden.

3.2.1 Taalontwikkeling

Om de taalontwikkeling van de peuters tussen drie en vier jaar te volgen is gebruik gemaakt van de passieve woordenschattoets van de Diagnostische Toets Tweetaligheid (DTT) bestaande uit 60 items, waarvan Nederlandse en Turkse parallelversies beschikbaar waren. Hoewel ontwikkeld voor kinderen van vier tot vijf jaar, bleek de toets in pilot-onderzoek bij driejarigen op de laagste prestatieniveaus (in de eigen taal!) nog voldoende te differentiëren. Op het tweede en derde meetmoment is tevens de Nederlandse en Turkse versie van de actieve woordenschattoets van de DTT gebruikt, die uit 40 items bestaat. Gebruik van deze toets op driejarige leeftijd, leek naar aanleiding van de pilot-studie niet raadzaam vanwege te sterke bodemeffecten.

Aangezien de Nederlandse en Turkse versies van de toetsen uit dezelfde items (woorden, concepten) bestonden en het vanwege te verwachten leereffecten niet raadzaam leek dezelfde items zowel in het Nederlands als in het Turks af te nemen, zijn voor de Turkse peuters de tests in tweeën gesplitst. Van de Nederlandse versie werden alleen de *oneven* (de *A-versie*) en van de Turkse alleen de *even* items afgenomen (de *B-versie*). Voor de vergelijking met de Turkse peuters zijn op dezelfde manier ook voor de Nederlandse peuters van elke toets twee subtests, een A- en B-versie, gecreëerd.

Op basis van de correlaties tussen de A- en B-versies zijn test-paralleltest en op basis van correlaties tussen twee afnames van dezelfde testversie op twee meetmomenten zijn test-hertest betrouwbaarheidscoëfficiënten bepaald. De betrouwbaarheidscoëfficiënt van de subtest Passieve Woordenschat, gemiddeld over versie, groep en meettijdstip, is .73; die van de subtest Actieve Woordenschat is .84. Beide zijn voldoende.

3.2.2 Algemene (nonverbale) cognitieve ontwikkeling

Met behulp van de nonverbaal af te nemen subtests Sorteren, Mozaïek en Combineren van de SON 2½-7 (Snijders & Snijders-Oomen, 1975) en de in het Turks vertaalde en afgenomen subtests Grootte en Classificeren van het toetspakket Ordenen (Van Kuyk & Kamphuis, 1991; Van Kuyk, 1994) is een indicatie van de algemene cognitieve ontwikkeling op driejarige leeftijd bepaald. De subtests van de SON geven te zamen een indicatie van de algemene (non-verbale) intelligentie. De subtest Grootte meet logo-mathematische begripskennis over omvang, hoeveelheid, aantal e.d., en kennis van comparatief en superlatief gebruik van deze begrippen; de gemeten vaardigheid wordt geacht een predictor te zijn van de latere reken- en wiskundige ontwikkeling. De subtest Classificeren meet semantisch-taxonomische begripskennis en wordt geacht een predictor van algemene ontwikkeling in verbale domeinen te zijn (met name taalbegrip en begrijpend lezen).

De betrouwbaarheden van de subtests van de SON zijn vanwege de eenmalige afname in dit onderzoek en de toepassing van de afbreekregel niet onderzocht. In het normerings- en standaardiseringsonderzoek bleken de hier ge-

bruikte subtests voldoende betrouwbaar (Snijders & Snijders-Oomen, 1975). De subtests van de toets Ordenen zijn wel op meer meettijdstippen afgenomen. De test-hertestbetrouwbaarheden van Grootte en Classificeren zijn gemiddeld respectievelijk .87 en .76 voor de Nederlandse en .90 en .82 voor de Turkse peuters. De correlaties tussen de vier subtests die gebruikt zijn om een index te maken van de algemene cognitieve ontwikkeling op driejarige leeftijd, zijn alle in de .40 tot .70 range; Cronbachs α van de index is .78 bij vier items.

3.2.3 Thuis taal en gezinskenmerken

In het interview met de ouder die het meest verantwoordelijk was voor de dagelijkse opvang en verzorging van het doelkind, in alle gezinnen de moeder, behalve in een waar de vader die rol heeft, werden verschillende zaken nagevraagd.

Thuis taal: In het interview werden zeven verschillende huiselijke situaties van mondeling taalgebruik geschetst, waaronder praten tijdens de maaltijd, praten als er familiebezoek is en praten om het kind iets uit te leggen. Gevraagd werd welke taal in deze situaties meestal en het meest werd gebruikt. Indien dit Nederlands was, werd de score 1 gegeven; indien dit Turks was werd het item als 0 gescoord. De zeven aldus gecodeerde items vormen een betrouwbare schaal: Cronbachs α is .91 in de volledige onderzoeksgroep en .94 in de Turkse subgroep. Op het tweede en derde meetmoment zijn deze vragen herhaald. De Cronbachs α 's zijn respectievelijk .93 en .88 in de volledige onderzoeksgroep. Op basis van deze schalen is de *longitudinale index* Thuis taal geconstrueerd, die weergeeft in welke mate er thuis Nederlands met het doelkind wordt gesproken.

Sociaal-economische status (SES): Gegevens over het opleidingsniveau van vader en moeder en het beroepsniveau van vader zijn gesommeerd tot een index van de sociaal-economische status. Gevraagd is naar het hoogste type onderwijs dat men heeft gevolgd en met een diploma heeft afgesloten. Uitgangspunt daarbij was een schaal die op het Nederlandse onderwijsstelsel was afgestemd. Ten behoeve van Turkse ouders die hun schoolloopbaan (grotendeels) in Turkije hebben doorlopen, was de schaal voor zover mogelijk opgebouwd uit equivalente schooltypen van

het Turkse onderwijsstelsel.

Ter bepaling van het beroepsniveau van de vader is gevraagd naar de mate waarin in het (ooit) uitgeoefende beroep met symbolische materialen en moderne (informatie)technologie moe(s)t worden omgegaan in plaats van met handgereedschap en zware, luidruchtige machines, waarin vergaderd moe(s)t worden met collega's, en stukken geschreven en cursussen gevolgd moe(s)ten worden om bij te blijven. Cronbachs α is .83 bij 10 items.

De genoemde indicatoren van de sociaal-economische status van het gezin zijn overigens *niet* herhaald gemeten. Verondersteld wordt dat er geen noemenswaardige veranderingen zijn opgetreden in sociaal-economische status tussen eerste en laatste meetmoment.

Frequentie Taalinteracties: In het interview is aan de ouders een aantal situaties voorgelegd van taalinteractie met het kind. Gevraagd is aan te geven hoe vaak (van meer malen per dag tot minder dan eens per maand) zij een dergelijke interactie hebben met het kind. Hiertoe zijn situaties geselecteerd die in theorie gekenmerkt worden door cognitief complexe en gedecontextualiseerde communicatie, waaronder met name praten om iets uit te leggen, praten over iets wat het kind heeft meegemaakt en (praten tijdens) lezen en voorlezen. Op het tweede en derde meetmoment zijn deze vragen herhaald. De resulterende schaal heeft op het eerste, tweede en derde meetmoment α 's van .80, .84 en .81 bij 12 items. Op basis hiervan is de *longitudinale index* Frequentie Taalinteracties gemaakt.

Intensiteit kinderopvang: Aan de ouders is gevraagd in maanden nauwkeurig, onderscheiden naar soort en aantal uren opvang per week, aan te geven in welke perioden ten behoeve van het doelkind tot dan toe van professionele kinderopvang buitenshuis gebruik is gemaakt. Op het tweede en derde meetmoment zijn deze vragen herhaald met betrekking tot de periode vanaf het vorige meetmoment. Op het derde meetmoment gaan de meeste kinderen al enige tijd (maximaal drie maanden) naar de basisschool; dit is gemakshalve beschouwd als een vorm van kinderopvang, met een intensiteit van zo'n 24 uur per week. Met deze gegevens is de *longitudinale index* Intensiteit Kinderopvang geconstrueerd die het produkt is van de totale duur en het gemiddelde aantal uren per week van opvang buitenshuis.

Tabel 1

De Nederlandse en Turkse taalontwikkeling van Turkse kinderen ($n = 27-31$), vergeleken met de Nederlandse taalontwikkeling van twee groepen Nederlandse kinderen (totaal $n = 77$); gemiddelden en standaardafwijkingen

Passieve woordenschat	leeftijd 3;0 jaar	leeftijd 3;6 jaar	leeftijd 4;0 jaar
Turkse peuters - Nederlandse versie	3.3 (9.0)	8.7 (17.2)	23.8 (21.6)
Turkse peuters - Turkse versie	43.6 (12.6)	57.0 (10.5)	62.9 (12.4)
Nederlandse laag-SES peuters	40.6 (17.9)	47.7 (18.2)	65.9 (14.5)
Nederlandse hoog-SES peuters	53.5 (18.9)	66.1 (16.5)	78.4 (11.8)
Actieve woordenschat			
Turkse peuters - Nederlandse versie		3.0 (8.2)	8.7 (13.6)
Turkse peuters - Turkse versie		34.2 (20.1)	47.8 (21.3)
Nederlandse laag-SES peuters		36.3 (17.4)	53.8 (18.6)
Nederlandse hoog-SES peuters		47.6 (17.6)	65.3 (15.2)

4 Resultaten

4.1 Woordenschatontwikkeling van Turkse en Nederlandse peuters

De geselecteerde items van de verschillende subtests zijn ongewogen opgeteld en gepercenteerd. In Tabel 1 zijn de gemiddelde scores op schalen van 0 tot 100 vermeld voor Turkse, Nederlandse laag-SES en Nederlandse hoog-SES peuters op drie (respectievelijk twee, bij de actieve woordenschat) meettijdstippen. Voor de Nederlandse groep zijn de A- en B-versies van de subtests Passieve Woordenschat en Actieve Woordenschat ten behoeve van de overzichtelijkheid samengenomen en gemiddeld.

De tabel toont dat de kennis van de Nederlandse taal van de driejarige Turkse peuters in dit onderzoek, aangenomen dat de passieve woordenschat daarvan een goede indicator is, *zeer gering* is (vanuit psychometrisch gezichtspunt is de meting, vanwege het bodem-effect, weinig zinvol). Het niveau ligt gemiddeld ongeveer tweeënhalve standaardafwijking (*sd*) lager dan dat van de Nederlandse laag-SES peuters en meer dan drie *sd* lager dan dat van de Nederlandse hoog-SES peuters. Dit komt nauw overeen met de resultaten van recent vergelijkend onderzoek naar de Nederlandse taalvaar-

digheid van Nederlandse en Turkse kleuters van vier en vijf jaar (Boogaard, et al., 1990; Verhoeven & Vermeer, 1992; De Jong, et al., 1995). Hoewel de Nederlandse taalkennis van de Turkse peuters na het tweede meetmoment, tussen drieneenhalf- en vierjarige leeftijd vrij sterk toeneemt, is het verschil met de Nederlandse laag-SES groep op het laatste meettijdstip, het moment van intrede in het basisonderwijs, niet noemenswaardig kleiner geworden.

De gegevens over de ontwikkeling van de actieve Nederlandse woordenschat vertonen ruwweg hetzelfde patroon: er is op drieënhaljarige leeftijd een verschil van ongeveer twee *sd* met de Nederlandse laag-SES groep en van tweeënhalve *sd* met de Nederlandse hoog-SES groep (ook hier geldt dat, psychometrisch gezien, door sterke bodem-effecten geen goede meting verricht is). Hoewel de actieve Nederlandse woordenschat van de Turkse peuters duidelijk toeneemt tussen het tweede en derde meetmoment, wordt de achterstand ten opzichte van beide Nederlandse groepen niet kleiner.

De kennis van het Turks van de Turkse peuters vergeleken met de kennis van het Nederlands van de Nederlandse peuters vertoont een ander ontwikkelingspatroon. Als groep scoren

Tabel 2

MANOVA voor herhaalde metingen, met A- en B-versies van de Passieve en Actieve Woordenschat-tests als afhankelijke variabelen, en met groep (Nederlands hoog-SES, Nederlands laag-SES en Turks) en meettijdstip (3, 6 en 4-jaar) als factoren ($n=108$)

2a: Passieve Woordenschat A: alleen Nederlandse versies

Variantiebron	SS	df	MS	F	p
Binnen cellen	86211.80	95	907.49		
Groep	166694.53	2	83347.26	91.84	.000
Binnen cellen	37451.26	190	197.11		
Meettijdstip	45372.32	2	22686.16	115.09	.000
Groep x Meettijdstip	1927.91	4	481.98	2.45	.045

2b: Passieve Woordenschat B: Nederlandse en Turkse parallel-versies

Variantiebron	SS	df	MS	F	p
Binnen cellen	69282.00	95	729.28		
Groep	11658.64	2	5829.32	7.99	.001
Binnen cellen	28318.71	190	149.05		
Meettijdstip	43623.02	2	21811.51	146.34	.000
Groep x Meettijdstip	1889.51	4	472.38	3.17	.015

2c: Actieve Woordenschat A: alleen Nederlandse versies

Variantiebron	SS	df	MS	F	p
Binnen cellen	43932.09	95	462.44		
Groep	62255.55	2	31127.78	67.31	.000
Binnen cellen	8764.69	95	92.26		
Meettijdstip	10205.42	1	10205.42	110.62	.000
Groep x Meettijdstip	2062.97	2	1031.48	11.18	.000

2d: Actieve Woordenschat B: Nederlandse en Turkse parallel-versies

Variantiebron	SS	df	MS	F	p
Binnen cellen	56930.58	95	599.27		
Groep	19464.33	2	9732.16	16.24	.000
Binnen cellen	9828.08	95	103.45		
Meettijdstip	12051.12	1	12051.12	116.49	.000
Groep x Meettijdstip	336.68	2	168.34	1.63	.202

de Turkse peuters in passieve kennis van de eigen taal op het eerste en tweede meetmoment zelfs iets hoger dan de Nederlandse laag-SES groep, ook al is de sociaal-economische positie van de Turkse groep in vergelijking met deze groep gemiddeld beduidend lager. Het verschil met de Nederlandse hoog-SES groep bedraagt een half tot één *sd*.

Tussen de passieve en actieve kennis die Turkse peuters hebben van het Turks en de kennis die zij hebben van het Nederlands bestaat op alle meetmomenten een groot verschil

dat varieert van een tot ruim twee *sd*'s. De verschillen zijn getest met *t*-toetsen voor afhankelijke steekproeven en zijn in alle gevallen significant. Wat betreft de passieve woordenschat wordt het verschil in kennis van het Turks en het Nederlands tussen drie- en vierjarige leeftijd duidelijk kleiner.

4.2 Hoofd- en interactie-effecten van 'meettijdstip' en 'groep'

Om de trends die uit Tabel 1 opdoemen op significantie te onderzoeken, zijn vier multiva-

riate variantieanalyses voor herhaalde metingen uitgevoerd met 'groep' en 'meettijdsp' als factoren en de scores op de A- en B-versies van de subtests als afhankelijke variabelen. In Tabel 2 zijn de resultaten weergegeven. Om na te gaan waardoor de effecten veroorzaakt worden, zijn verschillende afzonderlijke variantieanalyses uitgevoerd van de toetscores van alleen de Nederlandse laag-SES en Turkse groep, op alleen het eerste en tweede, en alleen het tweede en derde meettijdsp. De resultaten hiervan worden kortheidshalve niet in tabellen weergegeven, maar hierna in de tekst verwerkt. Vooraf moet nog opgemerkt worden dat de *power* van de toetsen om eerste orde interactie-effecten te ontdekken door de kleine *n* per subgroep tamelijk klein is; interactie-effecten moeten, in de terminologie van Cohen (1977), *sterk* zijn om ontdekt te kunnen worden.

In de analyses die in Tabel 2 worden gerapporteerd, blijken in alle gevallen significante hoofdeffecten gevonden te worden van de factoren 'groep' en 'meetmoment'. De hoofdeffecten van het meettijdsp wijzen erop dat over alle groepen de passieve en actieve woordenschat significant groeit tussen drie en vier jaar. De hoofdeffecten van groep bij de A-versies van de passieve en actieve woordenschat (alleen Nederlandse versies) worden veroorzaakt door significante verschillen tussen de beide Nederlandse groepen, en tussen de Turkse groep en beide Nederlandse groepen. Het hoofdeffect van de factor groep bij de B-versie van de passieve woordenschat (Nederlandse en Turkse parallelversies) wordt uitsluitend bepaald door de verschillen tussen de Nederlandse hoog-SES groep enerzijds en de Nederlandse laag-SES en Turkse groep anderzijds. In de analyses met alleen de Nederlandse laag-SES groep en de Turkse groep, waarbij verschillen in kennis van het Nederlands, respectievelijk in (parallele) kennis van het Turks worden getest, is er geen hoofdeffect van groep. Het hoofdeffect van de factor groep bij de B-versie van de actieve woordenschat (Nederlandse en Turkse parallelversie) wordt ook door significante verschillen tussen de Nederlandse laag-SES en de Turkse groep bepaald.

Er worden verschillende significante interactie-effecten van groep en meettijdsp gevonden. De interactie-effecten in de analyses van de A-versies van de passieve en actieve

woordenschat zijn toe te schrijven aan de geringere toename van de passieve Nederlandse woordenschat van de Turkse peuters vergeleken met de juist sterkere toename van de passieve Nederlandse woordenschat van de Nederlandse laag-SES peuters. De groei van de passieve Nederlandse woordenschat van de Turkse peuters tussen drie en drieënhalve jaar blijkt niet significant te zijn. De toename tussen drieënhalve en vier jaar is groter en wel significant, maar significant kleiner dan die van de Nederlandse laag-SES kinderen, die in deze fase gemiddeld sneller groeien. Het laatste blijkt ook te gelden voor de ontwikkeling van de actieve Nederlandse woordenschat in deze leeftijdsfase.

De passieve kennis van het Turks van Turkse peuters wijkt over het geheel genomen niet af van de passieve kennis die Nederlandse laag-SES kinderen hebben van het Nederlands. De actieve kennis van het Turks van Turkse kinderen is echter significant kleiner dan de actieve kennis van het Nederlands van Nederlandse laag-SES peuters.

De trend in de gemiddelde scores op de B-versie van de passieve woordenschattoetsen suggereert voorts een interactie-effect ten nadele van de Turkse taalontwikkeling van de Turkse peuters. Dit interactie-effect wordt in de variantie-analyse over alle groepen en meettijdsp, waarbij de vaardigheid in de eigen taal tussen de drie groepen wordt vergeleken, inderdaad gevonden (zie Tabel 2b). Bij nadere beschouwing blijkt het veroorzaakt te worden door een significant verschil in groei van de woordkennis tussen de Turkse en Nederlandse laag-SES peuters tussen drieënhalve en vier jaar. Dat er geen significant verschil in groei is tussen de Turkse peuters en Nederlandse hoog-SES leeftijdsgenootjes, bevreemdt in dit verband. Een mogelijke verklaring is dat er een plafondeffect is opgetreden bij de hoogst presterende kinderen die sterk oververtegenwoordigd zijn in de hoog-SES groep. Een andere verklaring is dat de *power* van de toets om het verwachte effect statistisch te ontdekken, te klein is.

Er wordt geen corresponderend interactie-effect bij de B-versie van de actieve woordenschattoets gevonden (zie Tabel 2d). Ook hier moet op de geringe *power* gewezen worden.

De kennis van Turkse peuters van het Ne-

derlands blijft in de loop van het vierde levensjaar sterk achter bij die van Nederlandse peuters. Hoewel er met name tussen drieënhalve en vier jaar een relatief sterke groei van de Nederlandse passieve en (in veel mindere mate) actieve woordenschat optreedt, is deze bij intree in de basisschool nog steeds veel kleiner dan die van Nederlandse leeftijdgenootjes. Combineren we deze bevinding met het eerder besproken gegeven dat de taalontwikkeling van de Turkse peuters in het Turks eerder trager dan sneller lijkt te verlopen, dan zou dit een aanwijzing kunnen zijn voor 'subtractieve tweetaligheid', zoals in de inleiding van dit artikel omschreven.

4.3 Tweetalige ontwikkeling van Turkse peuters in relatie tot achtergrondkenmerken

In Tabel 3 zijn de correlaties weergegeven van de scores van de Turkse peuters op de A- en B-versies van de passieve en actieve woorden-

schattoetsen met de algemene cognitieve ontwikkeling op driejarige leeftijd en de gezinskenmerken. De steekproefgrootte is klein (n -listwise=27; n -pairwise=27-31) en dat draagt het risico in zich dat het patroon van correlaties door enkele uitbijters wordt bepaald. Opvallende uitbijters zijn er overigens niet. Een tweede waarschuwing is dat de tabel correlaties weergeeft, geen causale relaties.

Desalniettemin lijken een paar conclusies gewettigd. De resultaten suggereren in de eerste plaats dat de determinering door sociale en individueel-cognitieve kenmerken van de taalontwikkeling van Turkse kinderen in het Turks fundamenteel afwijkt van de taalontwikkeling in het Nederlands. De Turkse taalontwikkeling van Turkse kinderen hangt veel sterker samen met hun algemene cognitieve ontwikkeling dan hun Nederlandse taalontwikkeling. Een mogelijke verklaring hiervoor is dat de index van de algemene cognitieve ontwikkeling op

Tabel 3

Correlaties van de subtestscores passieve en actieve Nederlandse en Turkse woordenschat met algemene cognitieve ontwikkeling en gezinskenmerken ($n=27-31$)

Leef- tijd kind	Subtestversie van de Diagnostische Toets Tweetaligheid	Algemeen Cognitief Niveau (3;0 jr)	SES van het gezin	Thuis- taal ¹	Freq. Taalinter- acties ¹	Intensiteit Kinder- opvang ¹
3;0	Passieve Woordenschat A (Nederlandse versie)	.29	.45 *	.27	.37 *	.51 **
	Passieve Woordenschat B (Turkse versie)	.45 *	.33	.24	.41 *	.34
3;6	Passieve Woordenschat A (Nederlandse versie)	.28	.72 ***	.55 ***	.44 *	.69 ***
	Passieve Woordenschat B (Turkse versie)	.62 ***	.15	-.09	.45 *	.24
4;0	Passieve Woordenschat A (Nederlandse versie)	.14	.70 ***	.48 *	.45 *	.36 *
	Passieve Woordenschat B (Turkse versie)	.65 ***	.60 ***	.22	.55 **	.02
3;6	Actieve Woordenschat A (Nederlandse versie)	.25	.67 ***	.64 ***	.46 *	.55 **
	Actieve Woordenschat B (Turkse versie)	.68 ***	.13	-.05	.34	.08
4;0	Actieve Woordenschat A (Nederlandse versie)	.22	.75 ***	.67 ***	.64 ***	.37 *
	Actieve Woordenschat B (Turkse versie)	.61 ***	-.08	-.23	.25	-.20

* $p < .05$; ** $p < .01$; *** $p < .001$

1 Deze variabelen zijn longitudinaal gemeten. De variabele intensiteit kinderopvang op het derde meetmoment geeft een indicatie van het aantal maanden dat het kind naar de basisschool gaat; dit varieert tussen 0 en 3 maanden.

driejarige leeftijd ten dele bestaat uit in het Turks gemeten (logo-mathematische en taxonomische) begripskennis die in theorie taalonafhankelijk, maar in praktijk, doordat begripskennis mede wordt gevormd in talige interacties, nauw verwant is met de woordenschat. Overigens suggereert dit gegeven dat transfer van algemene kennis en vaardigheden die geleerd zijn in de ene taal, naar (de verwerking van) een andere taal kennelijk in deze leeftijdsfase niet spontaan (d.i. door autonoom leren en informatieverwerken door het kind) tot stand komt. We komen daar nog op terug.

Omgekeerd hangt de Nederlandse taalontwikkeling van Turkse kinderen veel sterker samen met de sociaal-economische status van het gezin dan de Turkse taalontwikkeling. Een aannemelijke verklaring hiervoor is dat in hogere SES-gezinnen met de kinderen naar verhouding vaker Nederlands wordt gesproken (niet vermeld in de tabel is de correlatie tussen SES en Thuistaal, gemiddeld over de drie meetmomenten: $r=.68, p < .000$; Nota Bene: naarmate vaker Nederlands wordt gesproken, is de waarde van de variabele hoger). Dat hangt weer samen met een gemiddeld langere duur van het verblijf in Nederland en het feit dat door de ouders in Nederland onderwijs is gevolgd. Een verhoudingsgewijs groter aandeel van Turks in het taalaanbod van het gezin (kenmerkend voor het overgrote deel van de gezinnen) heeft echter geen vergelijkbare positieve invloed op de Turkse taalontwikkeling: de samenhang van de Thuistaal met de scores op de Turkstalige toetsversies is veel zwakker (en meestal niet significant).

De correlaties van de testcores op zowel de Turkse als Nederlandse toetsversies met de gerapporteerde regelmaat van 'hoog niveau' (gecontextualiseerde) taalinteracties in het gezin suggereren dat naast het aanbod in formele zin (aandeel van Nederlands versus Turks) ook de inhoudelijke kwaliteit van belang is. De correlaties van de Frequentie Taalinteracties wijzen er op dat de kwaliteit van het taalaanbod – of het nu Turks of Nederlands is – effecten van deze variabele op de taalontwikkeling modereert. In de gezinnen met een hogere sociaal-economische status wordt meer op hoog niveau met de kinderen gepraat, met name rond geschreven taal ($r=.62, p < .000$), en hoewel dit iets vaker lijkt te gebeuren in het Nederlands,

gebeurt dit in sommige gezinnen ook of alleen in het Turks. Daarom is het verband tussen Thuistaal en kwaliteit van het taalgebruik betrekkelijk zwak ($r=.35, p < .07$).

Ten slotte kan ook de intensiteit van het gebruik van (Nederlandstalige) kinderopvang een verklaring vormen voor het verband tussen SES en Nederlandse taalontwikkeling. De Intensiteit Kinderopvang hangt, niet onverwacht, significant samen met de Nederlandse taalontwikkeling van de Turkse peuters en niet (d.i. niet significant of negatief) met de ontwikkeling van de Turkse taalontwikkeling. Door gezinnen met een hogere sociaal-economische status wordt intensiever van kinderopvangvoorzieningen gebruik gemaakt, maar de correlatie is zwak (op het eerste meetmoment $r=.35, p < .07$). Overigens blijkt over het geheel genomen door de Turkse ouders in dit onderzoek slechts weinig gebruik gemaakt te worden van kinderopvangvoorzieningen, wat in overeenstemming is met gegevens uit hierboven aangehaald onderzoek. Op het derde meetmoment is de variabele Intensiteit Kinderopvang opgebouwd uit informatie over gebruik van voorschoolse opvang en over de tijd die het kind op het moment van testafname al naar de basisschool gaat (range 0-3 maanden). De richting van de correlaties is naar verwachting en suggereert een bevorderend effect op de Nederlandse taalontwikkeling, maar niet op de Turkse taalontwikkeling.

4.4 Zijn er aanwijzingen voor transfer?

De Tabellen 4 en 5 geven de correlaties voor de Turkse peuters tussen de scores op de A- en B-versies van respectievelijk de passieve en actieve woordenschattoetsen. De correlaties zijn nogal wisselend. Wat betreft de passieve woordenschat zijn de correlaties klein, maar overwegend positief (soms significant). Ook de cross-lag correlaties tussen de A- en B-versies over twee meetmomenten zijn overwegend positief. De trend is dat grotere passieve kennis van het Turks samengaat met grotere passieve kennis van het Nederlands en dat zou kunnen wijzen op cross-linguale transfer. Toch is dit gegeven vanwege inconsistenties in het patroon van correlatiecoëfficiënten moeilijk te evalueren. Het kan bovendien gaan om spuriëuze correlaties: de samenhangen worden veroorzaakt door andere variabelen, bijvoor-

Tabel 4

Correlaties tussen A- en B-versies van de subtests Passieve Woordenschat over drie meetmomenten voor de Turkse peuters (n=27-31); in cursief zijn de partiële correlaties weergegeven, controlerend voor de Frequentie Taalinteracties

Subtestversie en leeftijd kind	Passief-B Turks (3;0)	Passief-A Nederl. (3;6)	Passief-B Turks (3;6)	Passief-A Nederl. (4;0)	Passief-B Turks (4;0)
Passieve Nederlandse Woordenschat A (leeftijd 3;0)	.35 * .20	.77 *** .82 ***	.24 .06	.57 ** .50 **	.26 .11
Passieve Turkse Woordenschat B (leeftijd 3;0)		.38 * .07	.69 ** .67 **	.26 -.01	.58 ** .46 *
Passieve Nederlandse Woordenschat A (leeftijd 3;6)			.17 -.11	.60 ** .57 **	.43 * .18
Passieve Turkse Woordenschat B (leeftijd 3;6)				-.06 -.34 *	.64 ** .61 **
Passieve Nederlandse Woordenschat A (leeftijd 4;0)					.33 * -.03

* $p < .05$; ** $p < .01$; *** $p < .001$

beeld kenmerken van het taalgebruik in het gezin. Bij de actieve woordenschattoetsen vinden we overigens slechts een positieve samenhang tussen de Turkse en Nederlandse subtestscores (zie Tabel 5).

Om de invloed van gezinskenmerken uit te schakelen zijn tevens de partiële correlaties tussen de A- en B-versies van de verschillende subtests berekend, controlerend voor de Frequentie Taalinteracties als indicator van de kwaliteit van het taalgebruik in het gezin (deze variabele bleek eerder zowel met de Turkse als Nederlandse taalontwikkeling positief geassocieerd te zijn). In de Tabellen 4 en 5 zijn de partiële correlaties cursief gedrukt. De eerdere positieve verbanden bij de passieve woordenschat tussen de A- en B-versies blijken nu goeddeels te verdwijnen of negatief te worden, en bij de actieve woordenschat sterker negatief te worden (de cross-lag correlaties tussen sub-

tests van *dezelfde* versie, A of B, veranderen niet noemenswaardig). De huidige gegevens leveren, met andere woorden, geen duidelijke aanwijzing op voor positieve cross-linguale transfer. Dit betekent overigens niet dat er geen positieve transfer is. Wel veronderstellen we dat deze in deze leeftijdsfase en onder de beschreven condities van taalaanbod niet onafhankelijk is van gezinskenmerken die met name door de indicator van taalinteracties worden gerepresenteerd.

5 Discussie

Met de beperkingen van de hier gerapporteerde studie in gedachten, kunnen we concluderen dat de resultaten over het algemeen goed overeenkomen met eerdere onderzoeksbevindingen. Daaronder is het feit dat Turkse kinderen

Tabel 5

Correlaties tussen de A- en B-versies van de subtest Actieve Woordenschat over twee meetmomenten voor de Turkse peuters (n=27-31); in cursief zijn de partiële correlaties weergegeven, controlerend voor de Frequentie Taalinteracties

Subtestversie en leeftijd kind	Actief-B Turks (3;6)	Actief-A Nederlands (4;0)	Actief-B Turks (4;0)
Actieve Nederlandse Woordenschat A (leeftijd 3;6)	.20 .01	.75 *** .66 **	-.05 -.18
Actieve Turkse Woordenschat B (leeftijd 3;6)		.12 -.27	.77 *** .71 ***
Actieve Nederlandse Woordenschat A (leeftijd 4;0)			-.13 -.43 *

* $p < .05$; ** $p < .01$; *** $p < .001$

dominant in het Turks het onderwijs binnenkomen en op dat moment weinig tot zeer weinig vaardig zijn in het Nederlands.

De taalsituatie van de Turkse peuters in dit onderzoek wordt gekenmerkt door een overwegend Turkstalig aanbod, meestal van een dialect uit de streek van herkomst in Turkije. Naarmate het aandeel Nederlands in het taalaanbod groter was, waren de scores op de Nederlandse woordenschattoetsen duidelijk hoger. Een verhoudingsgewijs groter aandeel van het Turks in het taalaanbod bleek echter niet samen te hangen met de scores op de Turkse taaltoetsen. Als mogelijke verklaring hiervoor is het modererende effect van de kwaliteit van het taalgebruik genoemd. De kwaliteit van het taalgebruik hangt vooral samen met de sociaal-economische positie van de gezinnen.

De taalontwikkeling in het Turks tussen (ruwweg) de derde en vierde verjaardag van de Turkse peuters in dit onderzoek wijkt over het geheel genomen niet significant af van de taalontwikkeling van Nederlandse peuters uit de lagere sociale klasse. De verschillen in kennisniveau zijn klein (en soms ten gunste van de Turkse peuters); bij de actieve woordenschat zijn de verschillen overigens groter. We vonden echter dat tussen drieënhalf en vier jaar de groei van de passieve Turkse woordenschat van de Turkse peuters vertraagt, terwijl de ontwikkeling van de passieve kennis van het Nederlands van de Nederlandse laag-SES kinderen juist versnelt. Het interactie-effect is statistisch significant.

Een speculatieve verklaring luidt als volgt. Beide groepen peuters hebben gemeenschappelijk dat in deze fase intensiever van opvangvoorzieningen buitenshuis gebruik gemaakt wordt (inclusief de basisschool), waar Nederlands het taalaanbod domineert. Voor de Nederlandse laag-SES peuters heeft de verandering van het taalaanbod ten opzichte van thuis een gunstige invloed, omdat het taalaanbod in de opvangvoorzieningen ten opzichte van het taalaanbod thuis meerwaarde heeft (we vermoeden overigens dat dit *niet* of in veel mindere mate geldt voor Nederlandse hoog-SES peuters). Voor de Turkse peuters komt het Nederlandse taalaanbod in zekere zin te vroeg: weliswaar neemt hun kennis van het Nederlands rap toe, maar de ontwikkeling in de eigen taal lijkt eronder. Voor hen lijkt er in deze pe-

riode 'subtractieve tweetaligheid' op te treden.

Verschillen in groei van de eigen taal tussen Turkse en Nederlandse kinderen en tussen Turkse kinderen in Nederland en Turkije zijn ook in eerder onderzoek bij oudere kinderen vastgesteld. Onze gegevens suggereren dat het proces van subtractieve tweetaligheid tussen drieënhalf en vier jaar begint en onder andere door toenemend gebruik van de speelzaal en, uiteindelijk, de intree in de basisschool in gang wordt gezet. Om tot evenwichtige tweetalige ontwikkeling te komen lijkt de sterke toename van het aandeel van de Nederlandse taal in het taalaanbod van het merendeel van de Turkse peuters te vroeg te komen. We kunnen ons dan ook aansluiten bij pleidooien voor meertalig taalaanbod in kinderdagverblijven, peuterspeelzalen en kleutergroepen (zie voor een overzicht Veen & Vermeulen, 1993).

Tussen de Turkse en Nederlandse (passieve) woordenschat-ontwikkeling van de Turkse peuters bestaan zwakke positieve verbanden, wijzend op cross-linguale transfer van het Turks, dat al veel beter ontwikkeld is, naar de verwerving van het Nederlands. Voor *spontane*, door het kind zelf, *via* zelf geïnitieerde en bestuurd informatieverwerking, tot stand gebrachte transfer zijn weinig aanwijzingen gevonden. Dat hangt uiteraard samen met de ontwikkelingsfase en het feit dat, gemiddeld genomen, Nederlands in deze fase een marginale plaats heeft in het taalaanbod van Turkse peuters.

Uit een nadere beschouwing van de gegevens leiden we voorts de veronderstelling af dat als er transfer blijkt in deze fase, deze op een of andere manier *sociaal* - in taalinteracties - gemedieerd moet zijn. Nader onderzoek dat direct gericht is op transfer en op kenmerken van sociale interacties die cross-linguale transfer kunnen bevorderen, kan meer licht werpen op deze kwestie. Dergelijk onderzoek kan belangrijke informatie opleveren voor tweetalige taalstimuleringsprogramma's en voor speciale programma's voor tweede-taalonderwijs aan anderstalige kleuters. We kunnen veronderstellen dat het expliciet leggen van verbanden in didactische interacties tussen de eigen taal en het Nederlands op het niveau van algemene begrippen en metalinguïstische/taalanalytische kennis een belangrijk middel is om de effectiviteit van taalverwervingsprogramma's te bevorderen.

De beperkte omvang van de groep waarop de hier gerapporteerde gegevens over de tweetalige ontwikkeling van Turkse peuters gebaseerd zijn en de lage respons van Turkse gezinnen, staan niet toe dat groot gewicht aan de uitkomsten wordt toegekend. Hopelijk wordt er meer onderzoek geïnitieerd naar tweetalige ontwikkeling, zodat een preciezer en betrouwbaarder beeld kan ontstaan. Na deze voorbehouden gemaakt te hebben, durven we te stellen dat het in de inleiding geschetste theoretische model van tweetalige ontwikkeling in de context van het gezin door de gegevens ondersteund wordt. In dit verband is een belangrijke conclusie dat naast de spreiding van talen en taalvariëteiten in het taalaanbod, vooral de inhoudelijke kwaliteit van het taalgebruik (namelijk verhoudingsgewijs veel gedecontextualiseerde en cognitief-complexe ouder-kind communicatie) evenwichtige tweetalige ontwikkeling bevordert. De inhoudelijke kwaliteit is principieel onafhankelijk van de dominante taal of taalvariteit in het gezin *qua code*, maar hangt wel samen met de sociaal-economische positie van het gezin en de heersende taalgebruikstradities.

Literatuur

- Aarts, R., & Baaren, A. van (1990). De voorschoolse taalomgeving van Turkse kinderen. *Moer*, 7, 328-335.
- Appel, R., & Schaufeli, A. (1990). Meer Nederlands of meer Turks? *Forum der Letteren*, 31, 27-40.
- Berg, M.J.M. van den, & Vlug, I. (1993). *Bereik en kwaliteit kinderopvang voor 0 tot 4 jarigen uit achterstandssituaties*. Rotterdam: Rotterdams Instituut voor Sociologisch, Bestuurskundig en Milieukundig Onderzoek/Stedelijk Bureau FAO Rotterdam.
- Boeschoten, H.E., Verhoeven, L.T., & Vermeer, A. (1986). De taal van Turkse en Marokkaanse kinderen. In R. Appel (Red.), *Minderheden: taal en onderwijs*. Muiderberg: Couthino.
- Boogaard, M., Damhuis, R., Glopper, K. de, & Bergh, H. van den (1990). *De Nederlandse taalvaardigheid van allochtone en Nederlandse kleuters*. Lisse: Swets & Zeitlinger.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cummins, J. (1981). *Bilingualism and minority-language children*. Ontario: OISE Press.
- Cummins, J. (1984). Wanted: a theoretical frame for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 2-19). Clevedon: Multilingual Matters.
- De Temple, J.M., & Beals, D.E. (1991). Family talk: sources of support for the development of decontextualized language skills. *Journal of Research in Childhood Education*, 6, 11-19.
- Doets, C. (1992). Functionele geletterdheid onder volwassen Nederlanders. *Pedagogische Studiën*, 69, 388-399.
- Durán, R. (1987). Factors affecting development of second language literacy. In S.R. Goldman & H.T. Trueba (Eds.), *Becoming literate in English as a second language* (pp. 33-55). Norwood, NJ: Ablex.
- Extra, G., & Vallen, T. (1990). Meertaligheid en samenleving. In H.B. Entzinger & P.J.J. Stijnen (Red.), *Etnische minderheden in Nederland*. Meppel/Heerlen: Boom/Open Universiteit.
- Hakuta, K. (1986). *Mirror of language. The debate on bilingualism*. New York: Basic Books.
- Heath, S.B. (1983). *Ways with words*. Cambridge: Cambridge University Press.
- Heijden, H. van der (1992). De taalsituatie van Turkse kinderen in Nederland. In J. Rispen & B.F. van der Meulen (Red.), *Gezinsgerichte stimulering van kinderen in achterstandssituaties* (pp. 61-74). Lisse: Swets & Zeitlinger.
- Huls, E., & Mond, A. van de (1992). Some aspects of language attrition in Turkish families in the Netherlands. In W. Fase, K. Jaspaert & S. Kroon (Eds.), *Maintenance and loss of minority languages*. Amsterdam: John Benjamins.
- Jong, P.F. de, Klapwijk, M.J.G., & Leij, A. van der (1995). Cognitieve en sociaal-emotionele vaardigheden van kleuters in relatie tot hun etnische herkomst. *Pedagogische Studiën*, 72, 172-185.
- Kook, H. (1994). *Leren lezen en schrijven in een tweetalige context*. Amsterdam: IFOTT.
- Kuyk, J.J. van (1994). Does 'Ordering' lay the foundation of acquiring mathematics in young children? In F. Laevers (Ed.), *Defining and assessing quality in early childhood education* (pp. 197-212). Leuven: Leuven University Press.
- Kuyk, J.J. van, & Kamphuis, F.F.H. (1991). *Ordenen. Valideringsonderzoek van een observatieprogramma op het gebied van voorbereidend rekenen in de eerste groepen van de basisschool*. Arnhem: Centraal Instituut voor Toetsontwikkeling.

- Leseman, P. P. M. (1993). How parents provide young children with access to literacy. In L. Eldering & P. Leseman (Eds.), *Early intervention and culture* (pp. 149-171). Paris: Unesco Publishing.
- Leseman, P. (1994). Socio-cultural determinants of literacy development. In L. Verhoeven (Ed.), *Functional literacy: theoretical issues and educational implications* (pp. 163-184). Amsterdam/Philadelphia: John Benjamins.
- Leseman, P. P. M., Vergeer, M. M., Sijsling, F. F., Jap-A-Joe, S. R., & Şahin, S. (1992). *Amount, content and quality of parent-child interactions and the development of language and cognition*. Paper presented to the XXV International Congress of Psychology. Brussels, 21-25 July.
- Leseman, P. P. M., Sijsling, F. F., Jap-A-Joe, S. R., & Şahin, S. (1995). Gezinsdeterminanten van de cognitieve ontwikkeling van vierjarige Nederlandse, Surinaamse en Turkse kleuters. *Pedagogische Studiën*, 72, 186-205.
- Levelt, W. (1989). *On speaking*. Cambridge: Cambridge University Press.
- McDevitt, T. M., Hess, R. D., & Kashiwagi, K. (1987). Referential communication accuracy of mother-child pairs and later scholastic achievement: a follow-up study. *Merill-Palmer Quarterly*, 33, 171-185.
- Narrain, G., & Verhoeven, L. (1993). *Ontwikkeling van vroege tweetaligheid*. Tilburg: Tilburg University Press.
- Ogbu, J. (1992). Understanding cultural diversity and learning. *Educational Researcher*, 21, 5-14.
- Olson, S. L., Bayles, K., & Bates, J. E. (1986). Mother-child interaction and children's speech progress: a longitudinal study of the first two years. *Merill-Palmer Quarterly*, 32, 1-20.
- Snijders, J. T., & Snijders-Oomen, N. (1975). *Niet-verbale intelligentieschaal S.O.N. 2½-7*. Groningen: Wolters-Noordhoff.
- Snow, C. E. (1991). The theoretical basis for relationships between language and literacy development. *Journal of Research in Childhood Education*, 6, 5-10.
- Veen, A. M., & Vermeulen, A. C. A. M. (1993). *Kindercentra en meertaligheid*. Amsterdam: SCO Kohnstamm Instituut.
- Verhoeven, L. (1993). Literacy development in a multilingual context. In L. Eldering & P. Leseman (Eds.), *Early intervention and culture* (pp. 173-193). Paris: Unesco Publishing.
- Verhoeven, L. T., & Boeschoten, H. E. (1986). First language acquisition in a second language submersion environment. *Applied Linguistics*, 7, 241-256.
- Verhoeven, L. T., & Kuyk, J. J. van (1991). Peiling van conceptuele en metalinguïstische kennis bij de aanvang van het basisonderwijs. *Pedagogische Studiën*, 68, 415-425.
- Verhoeven, L., & Vermeer, A. (1992). Woordenschat van leerlingen in het Basis- en MLK-onderwijs. *Pedagogische Studiën*, 69, 218-234.
- Wells, G. (1987). Apprenticeship in literacy. *Interchange*, 18, 109-123.

Auteurs

P. P. M. Leseman is als hoofdonderzoeker verbonden aan het SCO Kohnstamm Instituut van de Faculteit der Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam.

S. Şahin studeert Turkologie in Leiden en was als assistent-onderzoeker verbonden aan het Rotterdams Instituut voor Sociologisch, Bestuurskundig en Milieukundig Onderzoek van de Erasmus Universiteit.

Correspondentie-adres: P. P. M. Leseman, Universiteit van Amsterdam, Faculteit POW, Wibautstraat 4, 1091 GM Amsterdam.

Abstract

Bilingual development of immigrant Turkish pre-schoolers in the Netherlands

P. P. M. Leseman & S. Şahin. *Pedagogische Studiën*, 1996, 73, 259-274.

In a small-scale longitudinal study, the development of active and receptive Dutch and Turkish vocabulary of 31 Turkish immigrant pre-school children was assessed from age three to four and compared with Dutch language development of 77 same-aged Dutch children. Using multivariate analysis of variance for repeated measurements, main-effects of 'group' (Turkish, Dutch working class and Dutch middle class) and 'age' were tested. Turkish pre-schoolers' development of Turkish language – for all of them the predominant home-language – appears

to be very similar to Dutch lower class children's language development. However, Turkish pre-schoolers' Dutch language development is rather far behind at every measurement point. Significant interaction-effects of group and age, furthermore suggest that in the period before and shortly after the fourth birthday the acceleration of Dutch language growth in Turkish children, due to increased use of pre-school and, from the fourth birthday onwards, primary school, is accompanied by a deceleration of Turkish language growth. A correlational analysis with home characteristics provides further clues as to the possible determinants of Dutch and Turkish language development of the Turkish pre-schoolers in this study.

Eigen-taalvaardigheid en schoolsucces van Turkse en Marokkaanse leerlingen*

R. Aarts, J. J. de Ruiter en L. Verhoeven

Samenvatting

In dit artikel wordt verslag gedaan van een onderzoek naar de relatie tussen eigen-taalvaardigheid en schoolsucces van Turkse en Marokkaanse leerlingen binnen het Nederlandse basisonderwijs. Het onderzoek vond plaats rond de overgang van basisonderwijs naar voortgezet onderwijs, omdat dit een kritisch moment is in de schoolloopbaan van leerlingen. Beslissingen die op dit moment genomen worden, spelen een rol in de verdere school- en beroeps carrière van leerlingen. Het is daarom van belang de factoren die invloed uitoefenen op deze keuzen alsmede de rol die de eigen taal daarbij kan vervullen, nauwgezet te onderzoeken. In het onderhavige onderzoek worden de prestaties van Turkse en Marokkaanse leerlingen op toetsen in de eigen taal en op de Eindtoets Basisonderwijs en de doorstroom naar het voortgezet onderwijs nader geanalyseerd. Daarbij wordt nagegaan wat de determinanten van schoolsucces van deze leerlingen zijn en welke rol de eigen taal daarbij speelt. De vaardigheid in de eigen taal wordt in dit artikel in verband gebracht met de vaardigheid in de tweede taal, het Nederlands, en de schoolcarrière van Turkse en Marokkaanse leerlingen. Er wordt geconcludeerd dat de vaardigheid in de etnische groepstaal een onderdeel van schoolsucces is, dat ook als zodanig gewaardeerd zou moeten worden.

1 Inleiding

Het schoolsucces van allochtone leerlingen op de Nederlandse basisschool wordt doorgaans eenzijdig geformuleerd in termen van de opbrengst in kennis en vaardigheden in het Nederlands als tweede taal. Daarbij wordt voorbijgegaan aan het feit dat kennis en vaardigheden in de etnische groepstaal evenzeer als belangrijke onderwijsleerresultaten kunnen worden gezien. Er is nog weinig bekend over de mate waarin de vaardigheid van allochtone leerlingen in de etnische groepstaal de predictie van later schoolsucces kan verbeteren. Het is mogelijk dat de beheersing van de eigen taal een belangrijke indicator vormt van het leervermogen van allochtone kinderen. Bovendien blijkt uit onderzoek dat kennis en vaardigheden die zijn opgedaan in de moedertaal van leerlingen overdraagbaar kunnen zijn naar overeenkomstige kennis en vaardigheden in de tweede taal. De door onderzoek gesteunde *afhankelijkheidshypothese* (Cummins, 1991; Verhoeven, 1994) geeft aan dat met name schriftelijke taalvaardigheden gemakkelijk overdraagbaar zijn van de ene taal naar de andere.

2 Tweetaligheid en schoolsucces

In het schooljaar 1991-1992 zat 7,6 procent van de 13-, 14- en 15-jarige Marokkaanse leerlingen en 4 procent van de Turkse leerlingen van deze leeftijd nog op de basisschool tegenover 0,3 procent van de Nederlandse leerlingen. In het voortgezet onderwijs bevond in 1991-1992 44 procent van de Turkse en 48 procent van de Marokkaanse leerlingen zich in het Lager Beroepsonderwijs (tegenwoordig VBO) tegenover 24 procent van de Nederlandse leerlingen. Van de leerlingen binnen het Algemeen Voortgezet Onderwijs ging 23 procent van de Turken en 18 procent van de Marokkanen naar het HAVO/VWO tegenover 46 procent van de Ne-

* Dit onderzoek is in opdracht van de Stichting voor Onderzoek van het Onderwijs (SVO) uitgevoerd in een samenwerkingsverband tussen het Instituut voor Toetsontwikkeling (Cito) en het werkverband Taal en Minderheden van de Katholieke Universiteit Brabant (KUB). Zie voor een verslag van het volledige onderzoek: Aarts R., Ruiter, J. J. de, & Verhoeven, L. (1993).

derlanders (Smeets, Martens & Veenman, 1995).

Uit het onderzoek van Mulder (1991) komt naar voren dat ongeveer de helft van de Turkse en Marokkaanse leerlingen een vervolgadvis krijgt dat gelijk is aan of lager dan LBO; slechts 7 procent krijgt een HAVO- of VWO-advies. Veel allochtone leerlingen komen op LBO- of MAVO-niveau terecht en ook binnen deze schooltypen kampen deze leerlingen met een achterstand. Ze halen prestaties die onder het gemiddelde van de klas liggen. De 'afstroom' van allochtonen is dan ook hoger dan die van autochtonen. Uiterwijk (1990, 1994) stelt vast dat allochtonen twee tot drie keer zoveel blijven zitten dan wel doorstromen naar schooltypen met een lager gemiddeld prestatieniveau in de tweede klas. Deze gegevens tonen aan dat Turkse en Marokkaanse leerlingen ten opzichte van Nederlandse leerlingen een achterstandspositie innemen in het onderwijs. In onderzoek wordt deze achterstand doorgaans aan twee oorzaken geweten: de taalproblematiek en de sociaal-economische status (SES). Gaandeweg de langere verblijfsduur van Turken en Marokkanen in Nederland lijkt de tweede genoemde oorzaak van zwaarder gewicht dan de eerste. Immers, de beheersing van het Nederlands verbetert, maar verbetering in maatschappelijke positie volgt niet in dezelfde mate. In het onderstaande zullen studies die uitspraken doen over deze problematiek besproken worden.

In het onderzoek van Kerkhoff (1988) naar de relatie tussen taalvaardigheid en schoolsucces zijn 179 leerlingen twee jaar lang in hun prestaties gevolgd. Onder deze groep bevonden zich 24 Marokkanen en 16 Turken. De prestaties van de Marokkanen en Turken werden vergeleken met andere allochtonen, dialectsprekers van het Nederlands en standaardsprekers van het Nederlands. Aan het einde van de basisschool werd 56 procent van de Marokkanen en 64 procent van de Turken een doorstroomadvies LBO gegeven en de overigen kregen een advies voor het AVO. Van de standaardsprekers van het Nederlands kreeg 36 procent een LBO-advies. De taalvaardigheid Nederlands bleek een minder grote invloed te hebben op het doorstroomadvies dan de SES. Kerkhoff concludeerde dat de verschillen tussen allochtone en autochtone leerlingen vooral

samenhangen met sociaal-economische kenmerken.

Van Langen en Jungbluth (1990) hebben de verschillen in onderwijspositie bij leerlingen in de derde klas van het voortgezet onderwijs onderzocht. Het onderzoek vond plaats in 1988 en behelsde een enquête onder allochtone en autochtone leerlingen op 107 scholen. Onder de allochtone leerlingen bevonden zich 93 Marokkanen en 122 Turken, van wie ongeveer drie kwart in het land van herkomst geboren was. De variantie in onderwijspositie in de gehele onderzoeksgroep werd vooral verklaard door de sociaal-economische positie van het gezin (deze verklaarde 17,4 procent van de variantie). De onderzoekers concluderen dat de sociaal-economische positie typerender is voor de onderwijspositie van allochtonen dan etniciteit. Echter, in nadere analyses per allochtone groep, bleken taalfactoren voor bepaalde groepen – en zeer zeker in het geval van de Marokkanen – toch een rol te spelen. 15,7 procent van de variantie in de onderwijspositie van Marokkanen werd verklaard door de spreektaal met de jongere broers en zussen. Voor de Turkse groep werd vastgesteld dat de factor 'moeilijkheid van het Nederlands' 11,6 procent van de variantie verklaarde.

Ook Driessen (1990) heeft de relatie onderzocht tussen de onderwijspositie van allochtone leerlingen en een aantal achtergrondkenmerken. De onderzoeksgroep bestond uit ruim 2000 leerlingen in groep 8 van de basisschool, waaronder 368 Turkse en 254 Marokkaanse leerlingen. Ook de vaardigheid in de eigen taal van allochtone leerlingen werd vastgesteld, maar er werd geen relevante samenhang tussen Nederlandse taalvaardigheid en eigen taalvaardigheid gevonden. Het voor de onderwijspositie meest relevante leerlingkenmerk bleek de onderwijservaring in Nederland te zijn. Wat gezinskenmerken betreft kwam het sociaal milieu als de belangrijkste factor naar voren. Rekening houdend met het sociale herkomstmilieu, bleek het effect van etnische herkomst afwezig bij rekenen en bij doorstroming naar het voortgezet onderwijs. Alleen bij taal werd een klein effect van etniciteit gevonden.

In een tweetal onderzoeken in het kader van de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid zijn effecten van etniciteit op schoolprestaties vastgesteld. Tesser en Mulder

(1990) gebruiken toetsscores als indicatoren voor schoolprestaties. Ook bij gelijke sociale positie blijken de reken- en taalscores van allochtone leerlingen lager te zijn dan die van autochtone leerlingen. Daarbij is er een sterk effect van verblijfsduur en thuistaal te constateren. Ze concluderen dat: "de achterstand van allochtone leerlingen niet zonder meer gereduceerd kan worden tot een vraagstuk van maatschappelijke positie, vergelijkbaar met de achterstandsproblematiek van Nederlandse arbeiderskinderen". In een latere publikatie (Mulder & Tesser, 1992) worden het effect van sociale kenmerken en het effect van etniciteit op schoolprestaties aan het einde van de basisschool, vervolgadvisie en schoolkeuze nader gespecificeerd. Bij prestaties blijft er na controle voor sociale-achtergrondkenmerken nog een vrij groot effect van etniciteit over. Bij vervolgadvisie en schoolkeuze daarentegen is het resterende effect van etniciteit na controle voor sociale-achtergrondkenmerken geheel verdwenen.

Van 't Hof en Dronkers (1992) maken in hun onderzoek onderscheid tussen generaties (de tussen-generatie en de tweede generatie) en tussen etnisch homogene en etnisch gemengde gezinnen. De onderzoeksgroep bestaat uit leerlingen die in 1989 de basisschool verlaten hebben, waaronder 247 Turkse en 250 Marokkaanse leerlingen. De onderzoekers stellen vast dat de sociaal-economische status de verschillen in rekenscores tussen de groepen geheel verklaart. De combinatie van SES en andere gezins- en leerlingkenmerken verklaart de verschillen in taalscores. Bij de informatiescore, het advies en de onderwijspositie blijft ook na controle voor de andere factoren het effect van de etnische factor aanwezig. Deze etnische factor blijkt echter een zeer verschillend effect te hebben voor de onderscheiden groepen. Voor Turkse leerlingen van de tussen-generatie werkt die nadelig: zij hebben een lagere onderwijspositie dan autochtone leerlingen met vergelijkbare prestaties, SES en andere leerling- en gezinskenmerken. Turkse leerlingen van de tweede generatie daarentegen bereiken na controle voor deze kenmerken een hogere onderwijspositie. Bij de Marokkanen heeft de tweede generatie zowel een hoger advies als een hogere onderwijspositie dan autochtone leerlingen met vergelijkbare kenmerken. De

auteurs concluderen dat de schoolloopbanen van de tweede generatie allochtone leerlingen, althans in het eerste leerjaar van het voortgezet onderwijs, zeker niet slechter zijn dan die van autochtone Nederlandse leerlingen met eenzelfde sociaal-economische achtergrond.

Uit een studie die uitgevoerd is in de bovenbouw van het voortgezet onderwijs (Hofman, 1993) komt naar voren dat de etnische herkomst, althans in deze fase van de onderwijsloopbaan, een belangrijke verklarende factor voor het schoolsucces van allochtone leerlingen is. Ook Fase (1994) is van mening dat bij de verklaring van de verschillen in onderwijspositie tussen allochtone en autochtone leerlingen de etnische factor niet verwaarloosd mag worden. In een vergelijkende studie naar het onderwijs aan etnische groepen in vijf Westeuropese landen stelt hij vast dat naast de sociaal-economische status en het migratieverschijnsel zelf (inpassingsproblemen die van korte duur zijn) ook etnisch gebonden hindernissen een rol spelen bij de schoolloopbaan van verschillende etnische groepen.¹

In verschillende onderzoeken wordt vastgesteld dat allochtone leerlingen doorgaans een hoger doorstroomadvies krijgen dan op grond van hun prestaties verwacht mag worden (onder andere Uiterwijk, 1994; Driessen, 1990; Mulder, 1991). Dit verschijnsel wordt wel aangeduid als *overadvisering*. Mogelijk is er sprake van twee soorten van overadvisering. De eerste is binnenschools; binnen een school worden aan allochtone leerlingen hogere adviezen gegeven dan aan autochtone leerlingen met dezelfde prestaties. Leerkrachten maken een inschatting van het onderwijsniveau dat een allochtone leerling aankan wanneer deze de Nederlandse taal beter zou beheersen. De leerkracht gaat er dan vanuit dat de taalachterstand in de loop van de tijd wel ingehaald zal worden. De overadvisering kan echter ook tussen scholen liggen. Allochtone leerlingen zitten vaak op scholen waar sprake is van een algemeen lager gemiddeld prestatieniveau. Het advies van de leerkracht is grotendeels gebaseerd op een vergelijking tussen de leerlingen onderling, zodat de hele 'adviseringsschaal' als het ware opschuift. Een betere leerling op zo'n school is mogelijk slechts een middelmatige leerling op een school met een hoger gemiddeld prestatieniveau. Zo ontstaan er tussen

scholen verschillen in de gemiddelde scores per doorstroomadvies. Uit het onderzoek van Mulder en Tesser (1992) komt naar voren dat de verklaring van de relatief hoge adviezen aan allochtone leerlingen zowel tussen als binnen scholen moet worden gezocht. De tussen- en binnenschoolse effecten op de adviezen aan individuele leerlingen zijn ongeveer even sterk.

Naast overadvisering wordt ook wel melding gemaakt van *overpositionering*. Allochtone leerlingen zouden doorgaans naar een hoger type van voortgezet onderwijs doorstromen dan autochtone leerlingen met hetzelfde prestatieniveau op de Eindtoets Basisonderwijs. Mulder (1991) stelt vast dat 9 procent van de Turkse leerlingen en 7,5 procent van de Marokkaanse leerlingen naar een schooltype gaat dat meer dan één niveau hoger ligt dan het geadviseerde schooltype. In het onderzoek van Mulder en Tesser (1992) is onderzocht welke factoren een rol spelen bij de 'overpositionering' van allochtone leerlingen. Het beroep van de vader en de taal- en rekenprestaties van de leerlingen blijken een significant effect te hebben op de schoolkeuze. Bij de Turkse en Marokkaanse leerlingen blijft er, ook na controle voor deze variabelen, een positief effect van herkomstland op schoolkeuze bestaan. Mulder en Tesser (1992) concluderen dat overpositionering bij allochtone leerlingen bestaat. Zij hebben echter geen antwoord gevonden op de vraag hoe het komt dat allochtone leerlingen vaker van het vervolgadvisie afwijken dan autochtone leerlingen met hetzelfde prestatieniveau en dezelfde sociale achtergrond.

Van 't Hof en Dronkers (1992) plaatsen kanttekeningen bij het hanteren van de termen overadvisering en overpositionering. Zij voeren nadere analyses op gegevens van schoolverlaters in 1989 uit, waarbij ze het etnische effect zowel bij advies als bij onderwijspositie tweemaal onderzoeken; eenmaal met en eenmaal zonder controle op milieu-, gezins- en leerlingkenmerken. Slechts als in beide gevallen het verschil met autochtone leerlingen significant is, accepteren ze de stelling dat er voor de betreffende groep sprake is van overadvisering en/of overpositionering. Onder deze condities blijkt er bij Turkse en Marokkaanse leerlingen geen sprake te zijn van overadvisering en alleen bij Marokkaanse leerlingen van overpositionering.

De resultaten en conclusies van de diverse onderzoeken blijken nogal uiteen te lopen. Dit heeft te maken met de moeilijkheden die kleven aan het vaststellen van onafhankelijke effecten van etnische herkomst en sociaal-economische status op het schoolsucces van leerlingen. Tesser (1993) noemt als bezwaar de onvergelykbaarheid van bepaalde opleidingscategorieën bij allochtonen en autochtonen en de onderlinge afhankelijkheid van de verdelingen van opleiding van de ouders en etnische herkomstgroepering. Naast de onafhankelijke factoren (SES, etniciteit) kunnen ook de afhankelijkste factoren op verschillende manieren vastgesteld worden (prestaties, rapportcijfers, doorstroom enzovoort). Bovendien is het van belang onderscheid te maken tussen etnische groepen en tussen generaties, omdat de verbanden voor deze groepen zeer verschillend kunnen zijn. Het trekken van algemene conclusies omtrent het effect van etnische herkomst en sociaal-economische status op het schoolsucces van allochtone leerlingen is dan ook niet mogelijk.

Uit het voorgaande is gebleken dat de onderwijspositie van allochtone leerlingen voor het grootste deel te verklaren is uit de sociaal-economische status van de ouders. Daarnaast lijkt ook een aantal taalspecifieke factoren een rol te spelen. Deze worden echter meestal in termen van achterstand gedefinieerd; men spreekt van een gebrekkige beheersing van het Nederlands. Daarbij wordt voorbijgegaan aan de vaardigheid in de etnische groepstalen. De vaardigheid in de eigen taal wordt in het algemeen niet als een onderdeel van schoolsucces gezien. Er is dan ook nog weinig zicht op de samenhang tussen vaardigheid in de etnische groepstaal en schoolsucces in het Nederlands. In het onderhavige onderzoek wordt deze samenhang nader onderzocht. We gaan daarbij uit van de volgende onderzoeksvragen:

- 1 Wat zijn de vaardigheden van Turkse en Marokkaanse leerlingen in hun eigen taal aan het einde van het basisonderwijs?
- 2 Welke factoren zijn van invloed op het schoolsucces van Turkse en Marokkaanse leerlingen bij de overgang van basisonderwijs naar voortgezet onderwijs?
- 3 Wat is de invloed van de eigen-taalvaardigheid van Turkse en Marokkaanse

leerlingen op hun schoolsucces in het Nederlands onderwijs?

3 Opzet van het onderzoek

3.1 Informanten

Het onderzoek vond plaats in 1991 en 1992. De onderzoeksgroep bestond uit 263 Turkse en 222 Marokkaanse leerlingen in groep 8 van de basisschool. De representativiteit van de onderzoeksgroep werd vastgesteld aan de hand van gegevens over de leerlingen, de sociaal-economische achtergrond van de gezinnen waaruit ze afkomstig waren en de geografische spreiding van de scholen. De gemiddelde leeftijd van de Turkse leerlingen was 12;08 jaar, van de Marokkaanse leerlingen 12;10 jaar. Drie kwart van de Turkse leerlingen en de helft van de Marokkaanse leerlingen was in Nederland geboren. Het merendeel van beide groepen (ca. 80 procent) heeft hun volledige schoolloopbaan in Nederland doorgebracht. De sociaal-economische status van de gezinnen waaruit de leerlingen afkomstig waren, was in het algemeen laag te noemen. De leerlingen uit de onderzoeksgroep zaten op 68 verschillende basisscholen verspreid over Nederland. De Turkse leerlingen hadden gemiddeld ruim 5 jaar OET gevolgd, de Marokkaanse leerlingen 4 jaar en 6 maanden. De gemiddelde lestijd voor OET was voor de Turkse leerlingen 120 minuten per week, voor de Marokkaanse leerlingen 113 minuten per week. Geen van de Turkse leerlingen in het onderzoek volgde buitenschools OET, van de Marokkaanse leerlingen deed een kwart dat wel.

De leerlingen maakten toetsen in de eigen taal en vulden vragenlijsten in. Ook door de groepsleerkracht en de OET-leerkracht werden vragenlijsten ingevuld. Daarnaast beschikten we over de resultaten die de leerlingen op de Eindtoets Basisonderwijs van het Cito behaald hadden. De informanten werden gevolgd in de eerste klas van het voortgezet onderwijs om de mate van schoolsucces vast te stellen. Ze bleken doorgestroomd te zijn naar 199 scholen voor voortgezet onderwijs. Van deze scholen waren er 153 bereid aan het vervolgonderzoek mee te werken, waardoor we gegevens konden verzamelen over 188 Turkse en 154 Marokkaanse leerlingen in het voortgezet onderwijs.

Naast de onderzoeksgroep in Nederland werden ook groepen leerlingen in Turkije en Marokko in het onderzoek betrokken. Dit waren 276 Turkse leerlingen van 6 verschillende scholen en 242 Marokkaanse leerlingen eveneens van 6 verschillende scholen. Deze groepen in de landen van herkomst werden in het onderzoek opgenomen om een vergelijking tussen hun taalvaardigheid en die van de leerlingen in Nederland mogelijk te maken.

3.2 Het toetsinstrumentarium

Het toetsinstrumentarium bestond uit de Toets Turkse Taal en Toets Arabische Taal, waarmee de vaardigheid in de etnische groepstaal vastgesteld werd. Deze toetsen werden in het kader van dit project geconstrueerd en zijn vanaf schooljaar 1992-1993 facultatieve onderdelen van de Eindtoets Basisonderwijs van het Cito. De toetsen bevatten taken op het gebied van technisch lezen, woordenschat, spelling, zinsbouw en begrijpend lezen. Voor een uitgebreide beschrijving van de constructie, inhoud en validiteit van deze toetsen verwijzen we naar Aarts et al. (1993). Om het schoolsucces in het Nederlands vast te stellen werden de resultaten op de Eindtoets Basisonderwijs van het Cito en het vervolgadvis van de leerkracht gebruikt. Zowel de Eindtoets als het vervolgadvis speelt een belangrijke rol bij de keuze voor een vorm van voortgezet onderwijs. Doorstroming naar het voortgezet onderwijs werd vastgesteld door aan de leerkrachten van de basisschool te vragen naar welke school voor voortgezet onderwijs de leerling doorgestroomd was. De mentor in het voortgezet onderwijs werd gevraagd naar de prestaties van de leerling in de eerste klas en de verwachtingen.

Tabel 1
*Variabelen met betrekking tot schoolsucces/
doorstroming en de instrumenten*

variabele	instrument
schoolsucces in BO	CITO Eindtoets BO Toets Turkse Taal/Toets Arabische Taal
vervolgadvis basisschool	vragenlijst groepsleerkracht
schoolsucces in VO: schooltype klas 1	vragenlijst mentor
doorstroming in VO: schooltype klas 2	vragenlijst mentor

gen voor de tweede klas (eventuele afstroom, doubleren). Tabel 1 geeft een overzicht van de afhankelijke variabelen en instrumenten waarmee schoolsucces in dit onderzoek vastgesteld wordt.

Als onafhankelijke variabelen – de mogelijke determinanten van schoolsucces – gelden de achtergrondkenmerken van de leerling, het gezin en de school. De benodigde gegevens werden verzameld door middel van vragenlijsten aan leerling, leerkracht en directie van de scholen. In de literatuur belangrijk geachte kenmerken (zie o.a. Kerkhoff, 1988; Van Langen & Jungbluth, 1990; Driessen, 1990) werden opgenomen in de vragenlijsten. Voor elk kenmerk werd een afweging gemaakt aan wie de vragen het beste gesteld konden worden. Tabel 2 geeft een overzicht van de achtergrondkenmerken en vragenlijsten.

3.3 Data-analyse

Allereerst is voor alle gehanteerde meetinstrumenten de interne consistentie gecontroleerd door gemiddelde item-totaalscore-correlaties en Cronbachs alpha's te berekenen. De alfa-betrouwbaarheid van de onderdelen van de Toets Turkse Taal varieerde van .69 tot .88. Die van de Toets Arabische Taal varieerde van .71 tot .97. Voor de Turkse leerlingen varieerde de alfa-betrouwbaarheid van de geconstrueerde schalen om achtergrondkenmerken vast te stellen van .67 tot .94 en voor de Marokkaanse leerlingen van .58 tot .94. De item-totaalcorre-

laties zijn in alle gevallen groter dan .25, hetgeen aangeeft dat de interne consistentie van de schalen redelijk is.

Om het niveau van vaardigheid in de etnische groepstalen van de leerlingen in Nederland te kunnen waarderen zijn hun scores op de onderscheiden toetstaken vergeleken met die van groepen leerlingen uit Turkije en Marokko. Met behulp van variantie-analyse zijn de verschillen tussen leerlingen in de beide landen op significantie getoetst.

De relatie tussen tweetaligheid en schoolsucces is op verschillende wijzen onderzocht. Allereerst is onderzocht vanuit welke leerling-, gezins- en schoolkenmerken de variatie op (deel)Eindtoetscores kan worden verklaard. Vervolgens is nagegaan in welke mate het advies voor vervolgonderwijs van de klasseleerkracht (mede) vanuit de beheersing van de etnische groepstaal kan worden voorspeld. Voorts is nagegaan in hoeverre die beheersing een additionele predictorvariabele vormt wanneer de feitelijke doorstroming in het voortgezet onderwijs als afhankelijke variabele wordt gekozen. Daarbij is gebruik gemaakt van zowel regressie-analyses als discriminantanalyses (dit laatste voor variabelen die niet op interval-niveau liggen). Bij deze analyses zijn we uitgegaan van de gegevens betreffende de onderinstromers en zijn de neveninstromers buiten beschouwing gelaten, omdat voor hen de verbanden anders en meer complex zijn. Van de onderzoeksgroep van Turkse leerlingen kun-

Tabel 2
Onafhankelijke variabelen en gebruikte instrumenten

variabelen		instrument: vragenlijst voor de
leerling- kenmerken	leeftijd, verblijfsduur, doublures	groepsleerkracht
	sociaal culturele oriëntatie	leerling
	zelfwaardering	leerling
	mate van OET	OET-leerkracht
gezins- kenmerken	sociaal-economische status	groepsleerkracht
	sociaal-culturele oriëntatie	leerling
	gezinscultuur: algemene stimulering specifieke stimulering schoolmotivatie capaciteit tot motiveren	OET-leerkracht, groepsleerkracht, leerling
	taalcontact	groepsleerkracht
	leerlingpopulatie	groepsleerkracht BO
school- kenmerken		

nen 218 leerlingen aangemerkt worden als onderinstromer, in de Marokkaanse groep zijn dat er 161.

4 Resultaten

4.1 Schoolsucces van Turkse leerlingen

De peiling van de taalvaardigheid in het Turks liet zien dat Turkse leerlingen aan het einde van de basisschool in het algemeen een hoge vaardigheid in de eigen taal hebben opgebouwd. De vergelijkingsgroep van leerlingen in Turkije scoort weliswaar op de meeste taken hoger, maar de verschillen zijn niet erg groot. Op het gebied van technisch lezen en zinsbouw zijn de verschillen gering. Op de onderdelen spelling en woordenschat zijn de verschillen significant. Op één onderdeel, begrip van lezen, scoren de leerlingen in Nederland zelfs hoger dan die in Turkije. Voor een uitgebreide beschrijving van de resultaten van de Turkse leerlingen op de Toets Turkse Taal verwijzen we naar Aarts et al. (1993) en Aarts (1994).

De Turkse leerlingen behalen op de Eindtoets Basisonderwijs een gemiddeld percentage goed van bijna 52 procent. Daarbij scoren ze op rekenen iets hoger, namelijk 55 procent, en op informatieverwerking lager, namelijk 49 procent. De Turkse leerlingen scoren ongeveer 20 procent lager dan het gemiddelde van de totale groep van deelnemers aan de Eindtoets Basisonderwijs in datzelfde jaar (1991). Bijna de helft van de Turkse leerlingen krijgt een advies voor het IBO of LBO, nog geen 15 procent krijgt een advies dat hoger ligt dan het MAVO. Vergelijken we deze cijfers met die van de Landelijke Evaluatie van het Onderwijsvoorrrangsbeleid (Mulder, 1991) dan blijkt er in de periode van 1989 tot 1991 weinig tot geen verbetering opgetreden te zijn. Mulder geeft de vervolgadvisen die 586 Turkse leerlingen in 1989 kregen. Ook van deze leerlingen kreeg nagenoeg de helft een vervolgadvis voor IBO of LBO, bijna een derde LBO-MAVO dan wel MAVO en 17 procent hoger dan MAVO.

Een groot deel van de Turkse leerlingen in ons onderzoek komt in een laag schooltype van voortgezet onderwijs terecht; 38 procent op een IBO- of LBO-school, een derde op LBO/MAVO of MAVO. 19 Procent van de leerlingen gaat naar MAVO/HAVO en 9% naar een

hoger schooltype (HAVO, HAVO/VWO). Wanneer we de feitelijke schooltypen waarin de Turkse leerlingen terecht komen vergelijken met het advies dat ze kregen, blijkt dat 48 procent een schooltype kiest dat hoger is dan het geadviseerde type. Tien procent kiest zelfs voor een schooltype dat twee niveaus hoger ligt. Dit is vergelijkbaar met de gegevens van Mulder (1991), die een percentage van 9 procent van de Turkse leerlingen noemt, die een schooltype kiezen dat meer dan 1 niveau hoger ligt dan het geadviseerde type. Bij de overgang naar de tweede klas gaat de overgrote meerderheid van de leerlingen door in hetzelfde schooltype, echter 9 procent stroomt af naar een lager type. 8 Procent van de leerlingen zal blijven zitten in de eerste klas. Uiterwijk (1990) noemt een afstroompercentage (doublure en lager schooltype) van 9 à 10 procent.

4.2 Schoolsucces van Marokkaanse leerlingen

De Marokkaanse leerlingen behalen redelijke resultaten in de mondelinge onderdelen van de Toets Arabische Taal, maar vrij lage resultaten in het schriftelijk Arabisch. Het verschil met de vergelijkingsgroep in Marokko is op alle onderdelen significant. Op het gebied van spelling en woordenschat zijn de verschillen tussen de leerlingen in Marokko en die in Nederland het grootst. De resultaten van de Marokkaanse leerlingen op de Toets Arabische Taal worden uitgebreid beschreven in Aarts et al. (1993).

Op de Eindtoets Basisonderwijs behalen de Marokkaanse leerlingen een gemiddeld percentage goed van bijna 49 procent. Daarbij scoren ze op rekenen het laagst, namelijk 47,5 procent, en op taal iets hoger, namelijk 50 procent. De Marokkaanse leerlingen scoren ongeveer 21 procent lager dan het gemiddelde van de totale groep van deelnemers aan de Eindtoets Basisonderwijs in datzelfde jaar. Bijna de helft van de Marokkaanse leerlingen krijgt een advies voor het IBO of LBO, nog geen 15 procent krijgt een advies dat hoger ligt dan het MAVO. In vergelijking met de gegevens van Mulder (1991) krijgen de Marokkaanse leerlingen in ons onderzoek iets lagere adviezen dan de 623 Marokkaanse leerlingen in de OVB-evaluatie die in 1989 in groep 8 van het basisonderwijs zaten. Van deze leerlingen kreeg 46 procent een vervolgadvis voor IBO of LBO, een derde LBO-MAVO dan wel MAVO en 17 procent

hoger dan MAVO.

Een groot deel van de Marokkaanse leerlingen in ons onderzoek komt in een laag schooltype van voortgezet onderwijs terecht; 40 procent op een IBO- of LBO-school, 37 procent op LBO/MAVO of MAVO. 13 Procent van de leerlingen gaat naar MAVO/HAVO en 8% naar een hoger schooltype (HAVO, HAVO/VWO). Wanneer we de feitelijke schooltypen waarin de Marokkaanse leerlingen terecht komen vergelijken met het advies dat ze gekregen hadden, blijkt 38 procent een schooltype te kiezen dat hoger is dan het geadviseerde type. 7,5 procent kiest zelfs voor een schooltype dat twee niveaus hoger ligt. Dit komt overeen met de resultaten van Mulder (1991), die eenzelfde percentage noemt. Een kwart van de Marokkaanse leerlingen kiest daarentegen voor een schooltype dat lager is dan het geadviseerde type. Bij de overgang naar de tweede klas gaat de overgrote meerderheid van de leerlingen door in hetzelfde schooltype, echter 8,7 procent stroomt af naar een lager type. 13 Procent van de leerlingen zal blijven zitten in de eerste klas. Dit cijfer is hoger dan de gegevens van Uiterwijk (1990), die een afstroompercentage (doublure en lager schooltype) van 9 à 10 procent noemt.

4.3 Determinanten van schoolsucces van Turkse leerlingen

De eigen-taalvaardigheid van Turkse leerlingen blijkt vooral verklaard te worden door de stimulering van de ouders, de zelfwaardering van de leerling en de mate waarin OET gevolgd is. Zie voor een uitgebreide beschrijving van deze resultaten Aarts et al. (1993) en Aarts (1994).

In Tabel 3 worden de achtergrondkenmerken, die een significante correlatie met één of meer factoren van schoolsucces in het Nederlands vertonen, weergegeven.

De verblijfsduur van Turkse leerlingen correleert licht negatief met vervolgadvis. Onderinstromers die langer in Nederland verblijven krijgen een lager advies. De leeftijd en het aantal doublures van de leerlingen correleert negatief met schoolsucces; oudere leerlingen en leerlingen die zijn blijven zitten (waarschijnlijk zijn dit in hoge mate dezelfde leerlingen), scoren significant lager op de Eindtoets Basisonderwijs, krijgen een lager vervolgadvis en stromen naar een lager schooltype van voortgezet onderwijs door. Leerlingen die zich sterker oriënteren op de Nederlandse taal en cultuur, scoren hoger op de Eindtoets Basisonderwijs, krijgen een hoger vervolgadvis en

Tabel 3

Pearson correlaties tussen achtergrondkenmerken en Eindtoetsscores (EB) van Turkse leerlingen; nonparr correlaties tussen achtergrondkenmerken en vervolgadvis en schooltype in klas 1 en klas 2 van Turkse onderinstromers (significantieniveau $p < .05$, ** $p < .01$)*

		EB	advies	schooltype klas 1	schooltype klas 2
leerling- kenmerken	verblijfsduur	-.11	-.12*	-.07	-.03
	leeftijd	-.22**	-.24**	-.20**	-.23**
	doublures	-.24**	-.30**	-.27**	-.25**
	sociaal-culturele oriëntatie	.21**	.18**	.12*	.18*
	zelfwaardering	.41**	.45**	.40**	.42**
gezins- kenmerken	sociaal-economische status	.21*	.25**	.28**	.30**
	sociaal-culturele oriëntatie	.15	.18*	.20**	.18
	algemene stimulering	.51**	.61**	.50**	.53**
	specifieke stimulering	.13	.18*	.17**	.14**
	schoolmotivatie	.40**	.40**	.47**	.33**
	capaciteit tot motiveren	.32**	.33**	.39**	.33**
school- kenmerken	taalcontact	.28**	.31**	.32**	.31**
	% allochtone leerlingen	-.14	.14*	.10	-.08
	% NL arbeiderskinderen	.05	-.03	-.18**	.11

stromen door naar hogere schooltypen. Verder blijkt sprake van een sterke samenhang tussen zelfwaardering en alle maten van schoolsucces.

De sociaal-economische status van het gezin correleert significant met alle schoolsuccesmaten. Voorts blijkt dat naarmate ouders zich meer op Nederland en de Nederlandse taal oriënteren, hun kinderen een hoger vervolgdadvies krijgen en doorstromen naar een hoger type van voortgezet onderwijs. De algemene stimulering door de ouders hangt sterk samen met alle schoolsuccesmaten. De stimulering die zich specifiek op scholing richt, correleert significant met het vervolgdadvies en met de doorstroming. De factoren schoolmotivatie en capaciteit van de ouders om het kind te motiveren, laten ook sterke verbanden met schoolsucces zien. Voor de factor taalcontact geldt hetzelfde: de mate van taalcontact in het Nederlands hangt samen met schoolsucces.

Het percentage allochtone leerlingen op school correleert licht positief met vervolgdadvies. Turkse leerlingen op scholen met een hoog percentage allochtone leerlingen krijgen dus een hoger vervolgdadvies dan die op scholen met een lager percentage allochtone leerlingen. Het percentage Nederlandse arbeidersleerlingen vertoont daarentegen een negatieve correlatie met de doorstroming naar klas 1 van het voortgezet onderwijs. Turkse leerlingen op scholen met een hoog percentage Nederlandse arbeidersleerlingen stromen door naar lagere types van voortgezet onderwijs.

De kenmerken van leerling, gezin en school die correleren met schoolsucces zijn in stapsgewijze regressie- en discriminant-analyses opgenomen om te bepalen welke van deze factoren het schoolsucces verklaren. De resultaten kunnen als volgt worden samengevat. De totaalscore op de Eindtoets Basisonderwijs van Turkse leerlingen wordt in sterke mate bepaald door de algemene stimulering vanuit het gezin. Daarnaast is ook de zelfwaardering van de leerling en het percentage allochtone leerlingen op school van belang. Samen verklaren deze drie variabelen 37% van de variantie in Eindtoetscores. Dezelfde drie factoren zijn eveneens van invloed op het vervolgdadvies. Daarnaast zijn de capaciteit tot motiveren en de sociaal-culturele oriëntatie van de ouders, de specifiek schoolgerichte stimulering door de ouders en

de mate van doubleren van de leerling van belang voor het vervolgdadvies. Deze factoren samen verklaren 53% van de variantie in vervolgdadvies.

De doorstroming naar de eerste klas van het voortgezet onderwijs wordt in hoge mate bepaald door de schoolmotivatie van de ouders en de zelfwaardering van de leerling. Daarnaast spelen ook het taalcontact in het gezin, de specifiek schoolgerichte stimulering door de ouders en de culturele oriëntatie van de leerling een rol. Het percentage verklaarde variantie dat door deze factoren wordt verklaard bedraagt 53%.

Het schooltype waarnaar de leerling in de tweede klas van het voortgezet onderwijs doorstroomt, wordt met name verklaard door de zelfwaardering van de leerling en de algemene stimulering door de ouders. Verder zijn van belang de schoolmotivatie van de ouders en de culturele oriëntatie van de leerling en van de ouders. Hier bedraagt het percentage verklaarde variantie 45%.

4.4 Determinanten van schoolsucces van Marokkaanse leerlingen

De eigen-taalvaardigheid van de Marokkaanse leerlingen wordt vooral verklaard door de mate waarin buitenschools OET is gevolgd, de schoolmotivatie van de ouders, de sociaal-culturele oriëntatie van de leerling en het percentage allochtone leerlingen op school. Voor een meer gedetailleerde beschrijving van deze resultaten verwijzen we naar Aarts et al. (1993).

In Tabel 4 worden de achtergrondkenmerken, die een significante correlatie met één of meer factoren van schoolsucces in het Nederlands vertonen, weergegeven.

De verblijfsduur van Marokkaanse leerlingen correleert licht positief met de score op de Eindtoets en met doorstroom. Naarmate Marokkaanse leerlingen langer in Nederland verblijven scoren ze hoger op de Eindtoets Basisonderwijs en stromen ze door naar hogere schooltypen. Oudere leerlingen krijgen een lager advies, maar stromen door naar hogere types voor voortgezet onderwijs. In de tweede klas van het voortgezet onderwijs, komen ze echter in lagere schooltypen terecht. Naarmate leerlingen meer op de Nederlandse taal en cultuur georiënteerd zijn behalen ze hogere scores

Tabel 4

Pearson correlaties tussen achtergrondkenmerken en Eindtoetsscores (EB) van Marokkaanse leerlingen; nonparr correlaties tussen achtergrondkenmerken en vervolgadvis en schooltype in klas 1 en klas 2 van Marokkaanse onderinstromers (* significantieniveau $p < .05$, ** $p < .01$)

		EB	advies	schooltype klas 1	schooltype klas 2
leerling- kenmerken	verblijfsduur	.18*	.13	.14*	.17*
	leeftijd	-.13	-.20*	.21**	-.23**
	sociaal-culturele oriëntatie	.23**	.16	.04	.10
	zelfwaardering	.23**	.35**	.25**	.27**
	buitenschools OET	.11	.32*	.27*	.28*
gezins- kenmerken	sociaal-economische status	.27*	.31*	.20	.26*
	algemene stimulering	.51**	.63**	.47**	.43**
	schoolmotivatie	.30*	.42**	.47**	.54**
	taalcontact	.23*	.37**	.28**	.17
school- kenmerken	% allochtone leerlingen	.11	.23*	.21*	.39**
	% NL arbeiderskinderen	-.07	-.18*	-.28**	-.23*

op de Eindtoets. Verder hangt de factor zelfwaardering van de leerling samen met alle schoolsuccesmaten. Marokkaanse leerlingen die veel naschools OET gevolgd hebben, krijgen een significant hoger vervolgadvis en stromen naar hogere schooltypen door.

Sociaal-economische status correleert significant met de score op de Eindtoets, met het vervolgadvis en met het schooltype in klas 2 van het voortgezet onderwijs. Algemene stimulering en schoolmotivatie van de ouders hangen sterk samen met alle schoolsuccesmaten. Taalcontact in het Nederlands hangt samen met de totaalscore op de Eindtoets, met het vervolgadvis en met de doorstroming naar het voortgezet onderwijs.

Het percentage allochtone leerlingen op school correleert significant met het vervolgadvis en met het bereikte schooltype in klas 1 en 2 van het voortgezet onderwijs. Het percentage Nederlandse arbeidersleerlingen op school hangt daarentegen negatief samen met advies en feitelijke doorstroming.

De kenmerken van leerling, gezin en school die correleren met schoolsucces zijn in stapsgewijze regressie- en discriminant-analyses opgenomen om te bepalen welke van deze factoren het schoolsucces verklaren. De resultaten van deze analyses kunnen als volgt worden getypeerd. De totaalscore op de Eindtoets Basisonderwijs van Marokkaanse leerlingen wordt met name verklaard door de algemene stimulering vanuit het gezin (26% verklaarde varian-

tie). Deze factor speelt samen met het gevolgde buitenschools OET ook een grote rol bij het vervolgadvis. Daarnaast zijn voor het vervolgadvis ook de schoolmotivatie van de ouders en de zelfwaardering van de leerling nog van belang. Samen verklaren deze factoren 48% van de variantie in vervolgadvis.

De algemene stimulering door de ouders en de zelfwaardering van de leerling zijn de verklarende factoren voor het schooltype dat de Marokkaanse leerlingen in klas 1 van het voortgezet onderwijs bereiken (29% verklaarde variantie). Beide factoren spelen ook een grote rol bij het schooltype dat de leerlingen in de tweede klas bereiken. Daarnaast is ook de schoolmotivatie van de ouders van belang. In dit geval bedraagt het totale percentage verklaarde variantie 59%.

4.5 De rol van de eigen-taalvaardigheid van Turkse leerlingen

Vervolgens willen we nagaan welke rol de eigen-taalvaardigheid van Turkse leerlingen speelt in hun schoolsucces in het Nederlandse onderwijs. Allereerst zijn daarom correlaties berekend tussen de eigen-taalvaardigheid en de verschillende schoolsuccesmaten. Daarna is de eigen-taalvaardigheid naast de scores op de Eindtoets Basisonderwijs als predictorvariabele opgenomen in discriminantanalyses ter verklaring van het vervolgadvis en de feitelijke doorstroming naar het voortgezet onderwijs. Bij deze analyses is de eigen-taalvaardig-

heid opgesplitst in een mondelinge en een schriftelijke factor.

De correlaties tussen de schoolsuccesmaten in het Nederlands en in het Turks worden weergegeven in Tabel 5. Er blijkt een duidelijk verband te bestaan tussen de scores op de mondelinge en schriftelijke onderdelen van de Toets Turkse Taal en het schoolsucces in het Nederlands onderwijs.

Tabel 5
Correlaties tussen schoolsuccesmaten in het Turks en het Nederlands

	Turks schriftelijk	Turks mondeling
Eindtoets Basisonderwijs	.40**	.36**
Vervolgadvies	.41**	.27**
Schooltype in klas 1 VO	.32**	.17**
Schooltype in klas 2 VO	.34**	.23**

In discriminantanalyses werden de verschillende toetsscores als predictorvariabelen opgenomen en het vervolgadvies respectievelijk het schooltype in klas 1 als afhankelijke variabele. Tabel 6 geeft de resultaten van deze discriminantanalyses weer.

Vooral de score op de deelttoets informatieverwerking van de Eindtoets Basisonderwijs blijkt een belangrijke predictorvariabele voor vervolgadvies te zijn. De score voor schriftelijke taalvaardigheid Turks vormt naast de (deel)scores op de Eindtoets Basisonderwijs, een additionele voorspeller van vervolgadvies. De feitelijke doorstroming naar de eerste klas van het voortgezet onderwijs wordt vooral verklaard door de scores op de onderdelen informatieverwerking en rekenen en in mindere mate door de score op het onderdeel Taal van de eindtoets.

Tabel 6
*Resultaten van discriminantanalyses met scores op de Eindtoets Basisonderwijs en scores op de Toets Turkse Taal als onafhankelijke en vervolgadvies respectievelijk schooltype in klas 1 als afhankelijke variabele (***) = significantieniveau $p < .001$*

	variabele	beta-coëfficiënt	canonische correlatie
Vervolgadvies	informatieverwerking	.60	
	taal	.29	
	rekenen	.21	
	Turks schriftelijk	.13	.70***
Schooltype klas 1	informatieverwerking	.57	
	rekenen	.40	
	taal	.19	.66***

4.6 De rol van de eigen-taalvaardigheid van Marokkaanse leerlingen

Voor de Marokkaanse leerlingen is op eenzelfde wijze als bij de Turkse leerlingen nagegaan welke rol de vaardigheid in de eigen taal speelt in hun schoolsucces in het Nederlands onderwijs. Ook hun resultaten op de Toets Arabische Taal zijn opgesplitst in een mondelinge en een schriftelijke factor. De mondelinge factor bestaat uit onderdelen in het Marokkaans-Arabisch, de schriftelijke factor uit onderdelen in het Standaard Arabisch. Tabel 7 geeft een overzicht van de correlaties tussen de schoolsuccesmaten in het Nederlands en de scores op de schriftelijke en mondelinge factor van de Toets Arabische Taal van de Marokkaanse leerlingen. De scores op de schriftelijke onderdelen Arabisch (Standaard Arabisch) blijken significant te correleren met het vervolgadvies en schooltype in klas 1 van het voortgezet onderwijs. De mondelinge vaardigheid in Arabisch (Marokkaans Arabisch) vertoont daarentegen een negatieve samenhang met het vervolgadvies.

Tabel 7
Correlaties tussen schoolsuccesmaten in het Arabisch en het Nederlands

	Arabisch schriftelijk	Arabisch mondeling
Eindtoets Basisonderwijs	.06	.10
Vervolgadvies	.30*	-.35*
Schooltype in klas 1 VO	.31*	.07
Schooltype in klas 2 VO	.15	.04

In een discriminantanalyse werden de verschillende toetsscores als predictorvariabelen opgenomen en het vervolgadvies als afhanke-

Tabel 8

Resultaten van discriminantanalyses met scores op de Eindtoets Basisonderwijs en scores op de Toets Arabische Taal als onafhankelijke en vervolgadvis respectievelijk schooltype in klas 1 als afhankelijke variabele (***) = significantieniveau $p < .001$)

	variabele	beta-coëfficiënt	canonische correlatie
Vervolgadvies	rekenen	.54	
	Arabisch mondeling	-.52	
	Arabisch schriftelijk	.44	
	informatieverwerking	.37	.65***
Schooltype klas 1	rekenen	.93	
	Arabisch schriftelijk	.37	.50***

lijke variabele. Tabel 8 geeft de resultaten van deze discriminantanalyse weer.

Met name de deelscore op het onderdeel rekenen van de Eindtoets blijkt een belangrijke predictorvariabele voor vervolgadvis. De score op de mondelinge onderdelen van de Toets Arabische Taal (negatief) en de score op de schriftelijke onderdelen van die toets en ten slotte de score op het onderdeel informatieverwerking van de Eindtoets zijn voorspellers van vervolgadvis. De schriftelijke vaardigheid in het Arabisch vormt dus voor Marokkaanse leerlingen – naast de scores op rekenen en informatieverwerking van de Eindtoets Basisonderwijs – een additionele voorspeller van vervolgadvis. Het bereikte schooltype in klas 1 van het voortgezet onderwijs wordt vooral voorspeld door de score op het onderdeel rekenen van de Eindtoets en de score op de schriftelijke factor van de Toets Arabische Taal. De toets in het Standaard Arabisch blijkt dus ook bij te dragen aan de voorspelling van de doorstroming naar het voortgezet onderwijs van Marokkaanse leerlingen. Opmerkelijk is ten slotte dat het onderdeel taal van de Eindtoets geen belangrijke rol bij advies, noch bij doorstroming speelt. Dit komt overeen met de bevinding van Kerkhoff (1988), die vaststelde dat scores op het vak rekenen van zwaarder gewicht op het doorstroomadvies waren dan de taalvaardigheid Nederlands.

5 Conclusie en discussie

5.1 Schoolsucces van Turkse en Marokkaanse leerlingen

De Turkse leerlingen in Nederland bereiken aan het eind van het basisonderwijs een taalvaardigheidsniveau in het Turks dat niet ver afwijkt van dat van leerlingen in Turkije. Dit in

weerwil van het feit dat zij slechts een beperkte hoeveelheid instructie in hun eigen taal krijgen. Voor de Marokkaanse leerlingen is het beeld anders. Zij beheersen het mondeling Marokkaans-Arabisch redelijk, maar hebben veel problemen met het Standaard Arabisch. Daarbij dient aangetekend te worden dat ook voor kinderen in Marokko de beheersing van de standaardtaal een lange leerweg vereist.

De Turkse leerlingen behalen op de Eindtoets Basisonderwijs een gemiddelde score die ongeveer 20 procent lager ligt dan het gemiddelde van de totale groep van deelnemers aan de Eindtoets Basisonderwijs in datzelfde jaar. Bijna de helft van de Turkse leerlingen krijgt een vervolgadvis voor het IBO of LBO, nog geen 15 procent krijgt een advies dat hoger ligt dan het MAVO. Bijna 40 procent van de Turkse leerlingen gaat naar een IBO- of LBO-school, een derde naar LBO/MAVO of MAVO. Toch kiest bijna de helft een schooltype dat hoger is dan het geadviseerde type, 10 procent kiest zelfs voor een schooltype dat twee niveaus hoger ligt. Bij de overgang naar de tweede klas stroomt 9 procent van de Turkse leerlingen af naar een lager schooltype. 8 Procent van de leerlingen blijft zitten in de eerste klas.

Ook de Marokkaanse leerlingen behalen op de Eindtoets Basisonderwijs een totaalscore die ongeveer 20 procent beneden het gemiddelde van de totale groep deelnemers ligt. Bijna de helft van de Marokkaanse leerlingen krijgt een vervolgadvis voor het IBO of LBO, nog geen 15 procent krijgt een advies dat hoger ligt dan het MAVO. 40 procent van de Marokkaanse leerlingen stroomt door naar een IBO- of LBO-school, 37 procent naar LBO/MAVO of MAVO. Toch kiest 38 procent van deze leerlingen een schooltype dat hoger is dan het geadviseerde type, 7,5 procent kiest zelfs voor

een schooltype dat twee niveaus hoger ligt. Een kwart van de Marokkaanse leerlingen kiest daarentegen voor een schooltype dat lager is dan het geadviseerde type. Bij de overgang naar de tweede klas stroomt 8,7 procent af naar een lager schooltype. 13 Procent van de leerlingen blijft zitten in de eerste klas.

5.2 Verklaring voor schoolsucces van Turkse en Marokkaanse leerlingen

Het onderzoek maakt duidelijk dat een groot aantal leerling-, gezins- en schoolvariabelen het schoolsucces van allochtone leerlingen in Nederland bepaalt. Wat betreft afzonderlijke variabelen vallen de volgende verbanden op.

Leerlingkenmerken

De verblijfsduur heeft voor Turkse leerlingen geen effect; voor Marokkaanse leerlingen werd er een positief effect van verblijfsduur op alle schoolsuccesmaten vastgesteld. Dit laatste gegeven komt overeen met de onderzoeken van Driessen, De Bot en Jungbluth (1989), De Jong (1985) en Tesser en Mulder (1990), waarin ook invloed van verblijfsduur op schoolsucces werd vastgesteld. We moeten hierbij aantekenen dat wij in onze analyses alleen de onderinstromers meegenomen hebben, waardoor de variatie in verblijfsduur niet zo groot is. De factor leeftijd heeft voor zowel Turkse als Marokkaanse leerlingen een negatief effect op schoolsucces. Oudere leerlingen presteren lager op de Eindtoets en stromen naar lagere schooltypes van voortgezet onderwijs door. De factor doubleren heeft voor Turkse leerlingen een negatief effect op schoolsucces; bij Marokkaanse leerlingen is er geen verband.

Naast de hiervoor genoemde globale achtergrondkenmerken van leerlingen blijken ook sociaal-culturele en motivationele aspecten van belang. Een positieve oriëntatie van de leerling op Nederland en de Nederlandse taal heeft zowel bij Turkse als Marokkaanse leerlingen een positief effect op schoolsucces. Deze bevinding komt overeen met de onderzoeken van Van Hout, Vallen en Stijnen (1989), Vermeer (1988) en Boers, Van den Bosch en Verhoeven (1987). In deze onderzoeken wordt veelal gesproken van licht positieve verbanden tussen een oriëntatie op de Nederlandse cultuur en schoolsucces. In ons onderzoek zijn de verbanden wat sterker. Ver-

der heeft zelfwaardering zowel bij Turkse als Marokkaanse leerlingen sterke verbanden met schoolsucces. Ook De Jong en Verkuyten (1989) stelden vast dat leerlingen met een hoge specifieke zelfwaardering de beste leerprestaties behaalden. De mate waarin naschools OET gevolgd is, hangt bij Marokkaanse leerlingen samen met vervolgadvisie en de bereikte schooltypes van voortgezet onderwijs.

Gezinskenmerken

De sociaal-economische status van het gezin hangt zowel bij Turkse als Marokkaanse leerlingen samen met schoolsucces. Dit verband komt in de meeste onderzoeken naar voren, o.a. in De Jong (1985), Van Langen en Jungbluth (1990) en Driessen (1990). Verder blijkt een culturele oriëntatie van de ouders op de Nederlandse taal en cultuur bij Turkse leerlingen samen te hangen met de score op rekenen, het vervolgadvisie en met het schooltype in klas 1.

De algemene stimulering vanuit het gezin en de mate waarin ouders hun kinderen motiveren voor school vertonen zowel bij Turkse als Marokkaanse leerlingen sterke verbanden met schoolsucces. De specifieke schoolgerichte stimulering hangt bij Turkse leerlingen samen met vervolgadvisie en met de bereikte schooltypes in het voortgezet onderwijs. De capaciteit tot motiveren van de ouders hangt bij Turkse leerlingen samen met alle schoolsuccesmaten. In het algemeen vinden wij veel sterkere samenhangen tussen de factoren van de gezinscultuur (ook wel onderwijsondersteunend thuisklimaat genoemd) dan in andere onderzoeken op dit terrein. In de onderzoeken van onder andere Tesser (1986) en Van Langen en Jungbluth (1990) werden slechts lichte samenhangen tussen gezinscultuur en schoolsucces vastgesteld.

Het taalcontact in het gezin hangt zowel bij Turkse als Marokkaanse leerlingen samen met schoolsucces. Dit gegeven komt overeen met de onderzoeken van Verhoeven (1989), Vermeer (1988), Driessen en De Bot (1990) en Tesser en Mulder (1990), die ook een verband tussen taalaanbod in het Nederlands en schoolsucces vaststelden.

Schoolkenmerken

Het percentage allochtone leerlingen op de basisschool hangt bij Turkse leerlingen samen

met het vervolgadvis; bij Marokkaanse leerlingen hangt het samen met vervolgadvis en met de bereikte schooltypes in het voortgezet onderwijs. Turkse en Marokkaanse leerlingen op scholen met een hoog percentage allochtone leerlingen krijgen dus een hoger vervolgadvis en Marokkaanse leerlingen stromen ook door naar hogere schooltypes van voortgezet onderwijs. Het lijkt erop dat we hier te maken hebben met tussenschoolse 'overadvisering' en 'overpositionering' zoals die ook in andere onderzoeken geconstateerd werden (o.a. Driessen, 1990; Uiterwijk, 1990; Mulder, 1991; Van 't Hof & Dronkers, 1992).

Het percentage Nederlandse arbeidersleerlingen op de basisschool hangt bij Turkse leerlingen negatief samen met het bereikte schooltype in klas 1 van het voortgezet onderwijs; bij Marokkaanse leerlingen zowel met vervolgadvis als de bereikte schooltypes. In het onderzoek van Tesser, Van der Werf en Mulder (1990) werd de invloed van deze factor op het vervolgadvis ook vastgesteld.

Samenvattend kunnen we stellen dat de volgende factoren het schoolsucces van Turkse leerlingen in hoge mate verklaren: de gezinscultuur en het taalcontact in het gezin, de zelfwaardering en sociaal-culturele oriëntatie van de leerling en het percentage allochtone leerlingen op school. Het schoolsucces van Marokkaanse leerlingen wordt vooral verklaard door de gezinscultuur, de zelfwaardering van de leerling en het gevolgde buitenschools OET.

5.3 De rol van de eigen taal

Voor de Turkse leerlingen werd een sterke samenhang tussen vaardigheid in het Turks en alle schoolsuccesmaten (Eindtoetscores, advies en doorstroming) vastgesteld. Voor de Marokkaanse leerlingen geldt dat er nauwelijks samenhang is tussen eigen taalvaardigheid en scores op de Eindtoets, maar wel met vervolgadvis en doorstroming naar het voortgezet onderwijs. De score op de schriftelijke factor van de Toets Turkse Taal vormt een additionele voorspeller van het vervolgadvis dat Turkse leerlingen krijgen. Voor Marokkaanse leerlingen blijkt de score op de schriftelijke factor van de Toets Arabische Taal een predictorvariabele voor zowel het vervolgadvis als het schooltype in klas 1 van het voortgezet onderwijs te zijn. Deze bevindingen staan in con-

trast met het onderzoek van Driessen (1990), die bij Turkse en Marokkaanse kinderen nauwelijks samenhang tussen eigen-taalvaardigheid en taalvaardigheid in het Nederlands vaststelde.

De scores op de toetsen in de eigen taal blijken extra informatie te kunnen geven over de vaardigheden van Turkse en Marokkaanse leerlingen. De vaardigheid in de etnische groepstaal van allochtone leerlingen zou als belangrijk onderdeel van schoolsucces opgevat moeten worden. Tevens kan de eigen-taalvaardigheid een voorspellende functie hebben inzake de doorstroming naar en het schoolsucces binnen het voortgezet onderwijs. Een belangrijke aanbeveling voor de onderwijspraktijk is dan ook het nader accentueren van de positie van etnische groepstalen binnen het schoolwerkplan. Dit houdt in dat aan het formuleren van eindtermen en van de leerweg om die eindtermen te kunnen bereiken de nodige aandacht wordt besteed. Daarnaast is het van belang om aan het einde van de basisschool de opbrengst van het onderwijs in de etnische groepstaal op zorgvuldige wijze te peilen. Daarbij kunnen de resultaten die in die taal zijn behaald worden meegewogen bij de advisering van vervolgonderwijs ten behoeve van allochtone leerlingen. Om een goede peiling van de vaardigheid van Turkse en Marokkaanse leerlingen aan het einde van de basisschool mogelijk te maken zijn de binnen dit onderzoek ontwikkelde toetsen ondergebracht bij het Centraal Instituut voor Toetsontwikkeling en kunnen ze als facultatief onderdeel bij de Eindtoets Basisonderwijs door scholen besteld worden.

In vervolgonderzoek kan de studie waarover hier wordt gerapporteerd op uiteenlopende wijzen worden uitgebouwd. Allereerst kan de studie worden uitgebreid naar andere allochtone groepen. Gegeven het specifieke acculturatieproces van de diverse etnische groepen in Nederland is een dergelijke uitbreiding zinvol te noemen. Verder verdient het aanbeveling om in aanvullend onderzoek groepen allochtone leerlingen over een langere periode in het voortgezet onderwijs te volgen, zodat beter zicht kan worden verkregen op de factoren die de doorstroming van allochtone leerlingen bepalen. Ten slotte lijkt het van belang om naast groots opgezette studies zoals de huidige, kleinschalige onderzoeken uit te voeren waarin

de schoolloopbaan van allochtone leerlingen in het voortgezet onderwijs wordt bestudeerd. Daarbij valt te denken aan gevalsstudies waarin specifieke leerlingkenmerken en interacties in gezin en school worden gerelateerd aan schoolsucces.

Noot

- 1 Fase (1994) verklaart de etnisch gebonden hindernissen van etnische groepen aan de hand van het begrip 'marktwaarde'. De marktwaarde van een groep bepaalt de relaties tussen de groep en de relevante omgeving. Drie terreinen zijn hierbij van belang, namelijk raciale verhoudingen in het maatschappelijk verkeer, etnische vooringenomenheid in het onderwijs en de mate van congruentie in opvoedingswaarden tussen gezin en school.

Literatuur

- Aarts, R., Ruiter, J.J. de, & Verhoeven, L. (1993). *Tweetaligheid en schoolsucces. Relevantie en opbrengst van etnische groepstalen in het basisonderwijs*. Tilburg: Tilburg University Press.
- Aarts, R. (1994). *Functionele geletterdheid van Turkse kinderen in Turkije en in Nederland*. De Lier: Academisch Boeken Centrum.
- Boers, M., Bosch, L. van den, & Verhoeven, L.Th. (1987). Taalaanbod en culturele oriëntatie als voorspellers van tweede-taalvaardigheid. *Tijdschrift voor Taalbeheersing*, 9, 166-180.
- Cummins, J. (1991). Conversational and academic language proficiency in bilingual contexts. *AILA Review, Reading in Two Languages*, 75-89.
- Driessen, G. (1990). De onderwijspositie van allochtone leerlingen. De rol van sociaal-economische en etnisch-culturele factoren, met speciale aandacht voor het Onderwijs in Eigen Taal en Cultuur. Nijmegen: ITS.
- Driessen, G., & Bot, K. de (1990). Leerlingkenmerken en taalvaardigheid in het Turks en het Nederlands. *Toegepaste taalwetenschap in artikelen*, 37, 83-91.
- Driessen, G., Bot, K. de, & Jungbluth, P. (1989). *De effectiviteit van het onderwijs in eigen taal en cultuur*. Nijmegen: ITS.
- Fase, W. (1994). *Ethnic divisions in Western European education*. Münster/New York: Waxmann.
- Hof, L. van 't, & Dronkers, J. (1992). *Onderwijsachterstanden van allochtonen: klasse, gezin of cultuur*. Amsterdam: SCO.
- Hofman, W.H.A. (1993). *Effectief onderwijs aan allochtone leerlingen. Een empirische studie naar de invloed van school- en klasfactoren op de loopbaan van allochtone leerlingen en autochtone leerlingen in het voortgezet onderwijs*. Delft: Eburon.
- Hout, R. van, Vallen, T., & Stijnen, S. (1989). Taaloordelen van allochtone en autochtone basisschoolleerlingen. *Pedagogische Studiën*, 66, 74-83.
- Jong, M.J. de (1985). Het onderwijsniveau van allochtone leerlingen in Rotterdam. In M.J. de Jong, *Allochtone kinderen op Nederlandse basisscholen; prestaties, problemen en houdingen* (pp. 3-12). Lisse: Swets & Zeitlinger.
- Jong, W. de, & Verkuyten, M. (1989). Culturele oriëntatie en leerprestaties Nederlandse taal van Turkse kinderen. *Pedagogische Studiën*, 66, 296-306.
- Kerkhoff, A. (1988). *Taalvaardigheid en schoolsucces; de relatie tussen taalvaardigheid Nederlands en schoolsucces van allochtone en autochtone leerlingen aan het einde van de basisschool*. Academisch proefschrift. Amsterdam/Lisse: Swets & Zeitlinger.
- Langen, A. van, & Jungbluth, P. (1990). *Onderwijskansen van migranten; de rol van sociaal-economische en culturele factoren*. Amsterdam/Lisse: Swets en Zeitlinger.
- Mulder, L. (1991). *De overgang van basis- naar voortgezet onderwijs van de o.v.b.-doelgroepen; advies, schoolkeuze en rapportcijfers in het eerste leerjaar*. Nijmegen: ITS.
- Mulder, L., & Tesser P. T. M. (1992). *De schoolkeuzen van allochtone leerlingen*. Nijmegen: ITS.
- Smeets, H. M. A. G., Martens, E. P., & Veenman, J. (1995). *Jaarboek minderheden 1995*. Houten/Zaventem: Bohn Stafleu Van Loghum.
- Tesser, P. (1986). *Sociale herkomst en schoolloopbanen in het voortgezet onderwijs*. Nijmegen: ITS.
- Tesser, P. T. M. (1993). *Rapportage minderheden 1993*. Rijswijk: Sociaal en Cultureel Planbureau.
- Tesser, P., & Mulder L. (1990). Concentratie en prestatie van allochtone leerlingen in het basisonderwijs. *Migrantenstudies*, 2, 31-44.
- Tesser, P., Werf G. van der, & Mulder L. (1990). Het onderwijsvoorrrangsbeleid en de onderwijskansen van leerlingen uit etnische minderheidsgroepen. In M. du Bois-Reymond & L. Eldering (red.), *Nieuwe oriëntaties op school en beroep; de rol*

van sexe en etniciteit (pp. 152-163). Amsterdam, Swets & Zeitlinger.

Uiterwijk, J. H. (1990). *Item- en testbias in de Eindtoets Basisonderwijs 1987*. Academisch proefschrift. Arnhem: CITO.

Uiterwijk, J. H. (1994). *De bruikbaarheid van de Eindtoets Basisonderwijs voor allochtone leerlingen*. Arnhem: Cito.

Verhoeven, L. (1989). Socio-cultural predictors of minority children's first and second language proficiency. In K. Deprez (Ed.), *Language and intergroup relations in Flanders and the Netherlands* (pp. 165-180). Dordrecht: Foris.

Verhoeven, L. (1994). Transfer in bilingual development: the linguistic interdependency hypothesis revisited. *Language Learning*, 44, 381-415.

Vermeer A. (1988). Socio-cultural factors and second language acquisition of Turkish and Moroccan children. In R. van Hout & U. Knoops (Eds.), *Language attitudes in the Dutch language area*. Dordrecht: Foris.

Auteurs

R. Aarts is als toegevoegd onderzoeker werkzaam bij het werkverband Taal en Minderheden van de Katholieke Universiteit Brabant.

J. J. de Ruiter is als toegevoegd onderzoeker werkzaam bij het werkverband Taal en Minderheden van de Katholieke Universiteit Brabant.

L. Verhoeven is werkzaam als hoogleraar Orthopedagogiek aan de Katholieke Universiteit Nijmegen.

Correspondentie-adres: Werkverband Taal en Minderheden, KUB, Postbus 90153, 5000 LE Tilburg.

Abstract

Language proficiency and School success of Turkish and Moroccan children

R. Aarts, J. J. de Ruiter & L. Verhoeven. *Pedagogische Studiën*, 1996, 73, 275-290.

This article gives a report of a study on the language proficiency of Turkish and Moroccan children at the end of primary school in the Netherlands. The research questions focus on the relationship between language proficiency and school success of Turkish and Moroccan children at the end of primary school and the beginning of secondary school. 263 Turkish and 222 Moroccan children who attended 68 different schools participated in the research. Among the informants tests scores in their native language and in Dutch, and data concerning the type of secondary school they visited were collected. In order to evaluate the children's first language proficiency, a reference group of children in Turkey and Morocco was also selected.

The results show that Turkish children attain a fairly high proficiency in Turkish, but Moroccan children have difficulties mastering Standard Arabic. Concerning their proficiency in Dutch and their success in the Dutch school system, Turkish and Moroccan pupils lag behind Dutch pupils. The family culture (the amount of stimulation from the family) and self-esteem of the pupil seem to be important factors influencing the school carrier of these pupils. Finally it was concluded that native language proficiency of Turkish and Moroccan children has a positive effect on their success in the Dutch school system and contributes to a better prediction of the carrier of these children in secondary school.

Samenvatting

Het onderzoek 'Taalpeiling yn Fryslân' (De Jong & Riemersma, 1994) is gericht op de beheersing van het Nederlands en het Fries door leerlingen aan het einde basisschool. De opzet van het onderzoek is gebaseerd op de taalpeiling van PPON, zodat een verantwoorde vergelijking van de resultaten met de resultaten van PPON gemaakt kan worden. De conclusie is, dat de leerlingen in Friesland aan het einde van de basisschool het Nederlands even goed beheersen als de andere leerlingen in Nederland. De beheersing van het Fries daarentegen blijft achter bij de in de kerndoelen geformuleerde verwachtingen. Het onderzoek besluit met de aanbeveling een doordacht taalbeleid voor tweetalig onderwijs in Friesland te ontwikkelen.

1 Inleiding

Sinds 1980 is Fries een verplicht onderdeel van het onderwijsaanbod op de basisscholen in Friesland. Uit een onderwijskundig perspectief ligt het voor de hand, dat de beleidsmaatregel tot invoering van het Fries als verplicht vak na een ruime invoeringsperiode geëvalueerd wordt. Geconstateerd kan worden, dat deze evaluatie niet heeft plaatsgevonden. Dit heeft als consequentie, dat de ontwikkeling van leerplannen, leermiddelen en de didactiek van het tweetalig onderwijs in Friesland gebaseerd is op veronderstellingen. Door het ontbreken van gegevens is er evenmin inzicht in de effecten van het tweetalig onderwijs op de beheersing van het Nederlands door zowel Friestalige als niet-Friestalige leerlingen. Dit betekent tevens, dat in de discussies met leerkrachten over de vormgeving van het tweetalig onderwijs en in de gesprekken met ouders over het belang ervan geen beroep gedaan kon worden op empirische gegevens. Om in deze behoefte aan empirische gegevens te voorzien is het onderzoek 'Taalpeiling yn Fryslân' (De Jong & Riemersma, 1994) uitgevoerd.

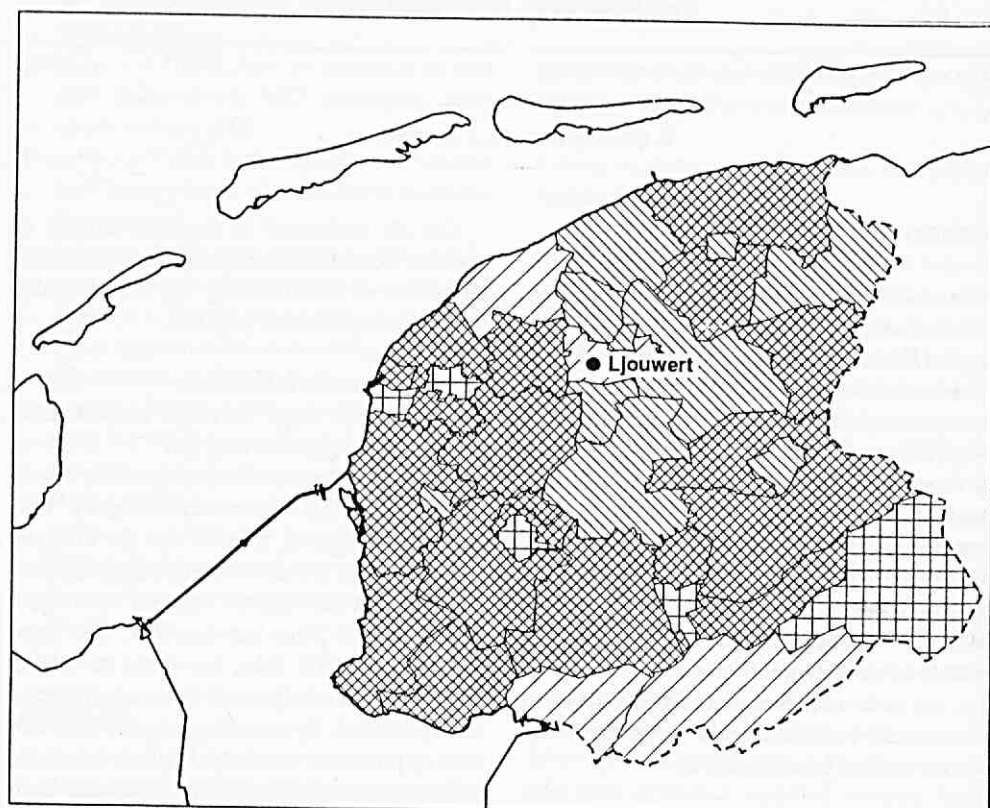
Om dit onderzoek in de juiste context te plaatsen wordt eerst een schets van de taalsituatie en van de ontwikkeling van het tweetalig onderwijs in Friesland gegeven.

1.1 De taalsituatie in Friesland

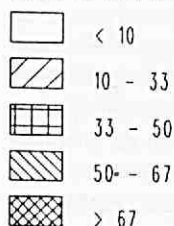
Friesland is de enige provincie in Nederland met een eigen (geschreven) taal.

Wat betreft het bereik en de spreiding van de taal kan verwezen worden naar de survey 'Taal yn Fryslân' (Gorter, Jelsma, Van der Plank & De Vos, 1984). Uit dit onderzoek blijkt, dat van de 600.000 inwoners van de provincie Friesland 94% Fries kan verstaan, 73% Fries kan spreken, 65% Fries kan lezen en slechts 10% Fries kan schrijven; 55% spreekt het Fries als moedertaal. De spreiding van de Friese taal over de provincie wordt duidelijk op het als Figuur 1 opgenomen kaartje dat gebaseerd is op deze survey¹.

Het kaartje laat een gevarieerd beeld zien: in de helft van de Friese gemeenten spreekt meer dan de helft van de inwoners thuis Fries, in Ooststellingwerf en in de kleinere steden spreken tussen één derde en de helft van de inwoners thuis Fries, terwijl in de gemeenten Het Bildt, Leeuwarden, Weststellingwerf en Lemsterland tussen de 10% en één derde van de inwoners thuis Fries spreekt. Op de Waddeneilanden wordt door minder dan 10% van de bevolking thuis Fries gesproken. Een dergelijke spreiding van de Friese taal heeft als gevolg dat het aandeel van Friestalige en niet-Friestalige leerlingen sterk varieert op de basisscholen in Friesland (Inspectie van het onderwijs 1989). Mede onder invloed van demografische ontwikkelingen is er de laatste decennia een sterke toename van gemengdtalige basisscholen in Friesland. Deze verschuiving wordt ook geconstateerd in 'Taal yn Fryslân' (Gorter et al., 1984: 15): *"in nivellearring tusken stêd en plattelân: de stêden binne mear en it plattelân minder 'Frysk' wurden."* ("een nivellering tussen stad en platteland: de steden zijn meer en



Frisian as first language (% of population)



— 5 km

Source: Gorfer et al. (1988) with adaptations

processed by means of Arc/Info (GIS)

Province Friesland, Databank/Cartography

Figuur 1. The language situation in Friesland (1980)

het platteland is minder 'Fries' geworden.”)

Het Fries functioneert niet alleen als de taal van thuis en in het onderwijs, maar ook in de verschillende domeinen. Er is sprake van een levendige Friese literatuur. Jaarlijks verschijnen er 80 tot 100 boeken, waarvan 20 kinderboeken. Er bestaat een succesvol professioneel theatergezelschap, 'Tryater', dat zowel het wereldrepertoire speelt als eigen produkties over het voetlicht brengt. Vrijwel elk dorp heeft een amateurtoneelvereniging die Friese stukken opvoert.

De regionale zender 'Omrop Fryslân' ver-

zorgt 53 uur per week Friese radio-uitzendingen en dagelijks een uur Friese televisie. In de beide regionale dagbladen worden dagelijks enige redactionele stukken in het Fries gepubliceerd, terwijl ongeveer 25% van alle familieberichten geheel of gedeeltelijk in het Fries gesteld zijn.

De Friese taal wordt voor hen die de provincie Friesland bezoeken, zichtbaar door de tweetalige plaatsnaamborden en Friese plaatsnamen op de bewegwijzering. Het gebruik van het Fries door de gemeenten en de provinciale overheid gaat echter verder: sinds de Tweede Wereldoorlog is de traditie gegroeid, dat het-

geen tijdens de vergaderingen van de gemeenteraad en van provinciale staten in het Fries gezegd wordt, in het Fries genotuleerd wordt. De burgers hebben het recht Fries te gebruiken in contact met de overheden, inclusief in contacten met de rechterlijke macht². Bovendien heeft het Fries een functie in het kerkelijk leven en in het sociale en economische leven.

1.2 Tweektalig basisonderwijs in Friesland

Tweektalig onderwijs in Friesland heeft een lange traditie. Na een schuchter begin als onderwijs buiten schooltijd tussen 1907 en 1937, werd Fries onder de noemer 'streektaal in levend gebruik' in 1937 in de wet opgenomen als een facultatief onderdeel van lezen en Nederlandse taal. Vanaf 1955 mag Fries gebruikt worden als instructietaal voor het aanleren van de basisvaardigheden lezen, schrijven en rekenen, terwijl vanaf 1980 Fries in de wet opgenomen is als een verplicht vak voor alle leerlingen op de basisscholen in Friesland.

Er is een relatie tussen de invoering van Fries als een verplicht vak en de verschuiving in motieven voor het onderwijs in het Fries. In de periode tot 1970 werd het onderwijs in het Fries onderbouwd met pedagogische motieven: de school had tot taak de Friestalige leerlingen in de moedertaal op te vangen met het doel het leren van het Nederlands te vergemakkelijken (transitie). Gaandeweg werd het culturele motief voor het onderwijs in het Fries meer beklemtoond. In het rapport van de Commissie Friese-Taalpolitiek (Ministerie van CRM, 1970) wordt aangegeven, dat juist de Friese taal in overheersende mate bepalend is voor het specifieke karakter van de Friese cultuur. Deze opvatting heeft uiteindelijk geleid tot wijziging van de Lager-Onderwijswet, waardoor het Fries met ingang van 1980 een verplicht vak is geworden op de basisscholen in Friesland.

Door de invoering van Fries als gewoon vak voor alle leerlingen op de basisscholen in Friesland is de algemene doelstelling van het tweektalig onderwijs in Friesland verschoven van het beheersen van de dominante taal (transitie) naar blijvende tweektaligheid. Terwijl vóór 1980 het onderwijs in het Fries vooral gericht was op de Friestalige leerlingen en met name bedoeld om deze leerlingen met behulp van het Fries goed Nederlands te leren, is dit onderwijs thans bedoeld om zowel de Friesta-

lige als de niet-Friestalige leerlingen beide schooltalen, Nederlands en Fries, goed te leren beheersen. Voor beide talen zijn identieke kerndoelen vastgesteld.

2 Vraagstelling en opzet van het onderzoek

Het onderzoek 'Taalpeiling yn Fryslân' is gericht op het verwerven en interpreteren van empirische gegevens teneinde een bijdrage te leveren tot de evaluatie van het tweektalig onderwijs in Friesland zoals aangegeven in de inleiding. Deze gegevens kunnen gebruikt worden om uitspraken te doen over de relatie tussen de kerndoelen Fries en Nederlands en de onderwijspraktijk in Friesland. Op grond van deze overwegingen zijn onder andere de volgende vraagstellingen geformuleerd:

- 1 Wat is het niveau van de taalvaardigheid Fries en Nederlands van zowel Friestalige als Nederlandstalige leerlingen aan het eind van de basisschool in Friesland?
- 2 In hoeverre is er sprake van overeenkomsten en verschillen tussen de gemeten taalvaardigheid Nederlands op de basisscholen in Friesland en de resultaten van het taalpeilingsonderzoek van PPON?

De opzet van het onderzoek 'Taalpeiling yn Fryslân' is sterk bepaald door de vormgeving van de eerste taalpeiling van de Periodieke Peiling van het OnderwijsNiveau (PPON) (Zwarts, 1990), die gericht is op de beheersing van de vier vaardigheden luisteren, lezen, spreken en schrijven. Voor het meten van het niveau van de beheersing van het Nederlands is een selectie gemaakt uit de toetsen die ontwikkeld zijn voor de eerste taalpeiling einde basisonderwijs. De projectleiding van PPON heeft ons de data van de gekozen toetsen beschikbaar gesteld. Hierdoor was het mogelijk de resultaten op de desbetreffende toetsen voor Nederlands van leerlingen in Friesland te vergelijken met de resultaten van de leerlingen in Nederland als geheel. Analooq aan deze toetsen voor Nederlands zijn voor Fries gelijkwaardige toetsen ontwikkeld; daarnaast is voor Fries een meerkeuze cloze-toets ontwikkeld. Naast globale toetsen voor de vier taalvaardigheden zijn deelvaardigheidstoetsen afgenomen voor tech-

nisch lezen en voor spelling. Bovendien heeft elke leerling van zichzelf aangegeven in hoeverre hij/zij Fries kan verstaan, lezen, spreken en schrijven (gerapporteerd taalgedrag).

Van de taalvaardigheden is in de eerste plaats het feitelijke beheersingsniveau aan het einde van de basisschool vastgesteld. De analyse van de gegevens heeft betrekking op de samenhang tussen de gemeten taalvaardigheid en twee groepen onafhankelijke variabelen. Als leerlingkenmerken zijn in de analyse betrokken de thuistaal van de leerling en het geslacht; als schoolkenmerken de taalomgeving van de school, de sociaal-economische status van de school, het gebruikte onderwijsleerpakket voor Nederlands en het onderwijsaanbod voor Fries (inclusief het gebruik van Fries als voertaal).

Aan het onderzoek hebben 48 (van de ruim 500) basisscholen in Friesland deelgenomen, ingedeeld naar het percentage Friestalige leerlingen:

- Categorie A: 10%-25% Friestalige leerlingen (8 scholen);
- Categorie B: 45%-55% Friestalige leerlingen (10 scholen);
- Categorie C: 75%-90% Friestalige leerlingen (16 scholen).

Het gemeenschappelijk kenmerk van deze scholen is, dat Fries tussen de 10% en 30% van de onderwijstijd als voertaal gebruikt wordt en dat gemiddeld drie kwartier per week onderwijs in het Fries als vak gegeven wordt. Daarnaast zijn twee categorieën scholen onderscheiden die gekenmerkt worden door een afwijkend onderwijsaanbod:

- Categorie A': 10%-25% Friestalige leerlingen; op deze scholen wordt Fries niet als voertaal gebruikt (9 scholen);
- Categorie C': 75%-90% Friestalige leerlingen; op deze scholen wordt twee maal zoveel tijd aan Fries als vak besteed als op de scholen in categorie C (5 scholen).

In verband met de celgrootte zijn alleen scholen met tenminste 10 leerlingen in groep 8 in het onderzoek betrokken. De selectie op basis van deze criteria heeft tot gevolg, dat van het onderzoek uitgesloten zijn zowel de scholen met heel weinig Friestalige leerlingen (bijvoorbeeld op de Friese waddeneilanden en in de

Stellingwerpen) als de (meestal kleine) scholen die het Fries als instructietaal gebruiken bij het aanleren van de basisvaardigheden lezen, schrijven en rekenen.

De schriftelijke toetsen zijn gemaakt door 788 leerlingen, te weten 426 Friestalige en 362 Nederlandstalige leerlingen. De mondelinge toetsen zijn gemaakt door een steekproef uit deze leerlingen.

3 Resultaten

In deze paragraaf wordt een overzicht gegeven van de toetsresultaten op de diverse deelvaardigheden. Zoals reeds in de inleiding is aangegeven, wordt de nadruk gelegd op de beheersing van het Nederlands door de leerlingen in Friesland in vergelijking met de leerlingen in Nederland als geheel.

De gemiddelde scores van zowel de Friestalige en Nederlandstalige leerlingen als van de leerlingen van de onderscheiden categorieën basisscholen worden vergeleken met de resultaten van de leerlingen die in het peilingsonderzoek van PPON de desbetreffende toets hebben gemaakt.

3.1 De beheersing van het Nederlands

3.1.1 Luistervaardigheid

Voor het toetsen van de luistervaardigheid Nederlands is gebruik gemaakt van een gedeelte van de toets 'Luisteren naar rapporterende teksten' van het PPON-onderzoek (PPON, 1991: 41-42). De toets bestaat uit de volgende onderdelen:

- 1 Vier open kennisvragen over een mededeling betreffende een afgelaste circusvoorstelling.
- 2 Twee meerkeuze-vragen en een open vraag betreffende een nieuwsbericht.
- 3 Het maken van een korte samenvatting over een informatiebericht, waar twee elementen in voor moeten komen.

De maximumscore is 9 punten.

Uit de variantie-analyse blijkt dat de variabelen thuistaal ($F=3.651$, $df=1,564$, $p=.057$) en taalomgeving ($F=.578$, $df=2,564$, $p=.562$) geen significante verschillen laten zien. Dit betekent, dat er geen verschil is tussen de resultaten

van de Friestalige leerlingen (7.23) en van de Nederlandstalige leerlingen (7.42): gemiddeld wordt in beide groepen 80% van de opgaven goed gemaakt. Ook de categorieën die qua taalomgeving het verst uit elkaar liggen, vertonen geen verschil: de Friestalige leerlingen in een overwegend Friestalige situatie (cat. C) en de Nederlandstalige leerlingen in een overwegend Nederlandstalige situatie (cat. A) scoren respectievelijk 7.30 en 7.25.

Door middel van variantie-analyse is eveneens nagegaan of de mate van het gebruik van Fries als voertaal en de tijdsinvestering voor Fries samenhangen met de resultaten van de leerlingen op het gebied van de luistervaardigheid Nederlands. Het blijkt dat er geen sprake is van samenhang tussen de resultaten op de toets luistervaardigheid Nederlands en de mate van het gebruik van het Fries als voertaal ($F=1.70$, $df=1,285$, $p=.681$). De leerlingen op scholen waar het Fries *niet* als voertaal gebruikt wordt (cat. A') ($n=130$), scoren met een gemiddelde van 7.31 dus *niet* significant hoger op het gebied van de luistervaardigheid Nederlands dan de leerlingen van basisscholen waar het Fries *wel* als voertaal wordt gebruikt (cat. A) ($n=159$), die een gemiddelde van 7.25 behalen. Tevens blijkt dat er geen sprake is van samenhang met de tijdsinvestering voor het vak Fries. Friestalige leerlingen op scholen waar per week tweemaal zoveel tijd aan het vak Fries besteed wordt (cat. C') ($n=50$), scoren met een gemiddelde van 7.18 *niet* significant lager dan Friestalige leerlingen op scholen in een vergelijkbare taalsituatie met een beperktere tijdsinvestering voor het vak Fries (cat. C) ($n=185$), die tot een gemiddelde score van 7.19 komen ($F=.005$, $df=1,233$, $p=.946$).

Er blijken dus geen significante verschillen te zijn tussen de Friestalige en Nederlandstalige leerlingen van de scholen van de diverse categorieën. Als we de scores nu vergelijken met het PPON-gemiddelde, dan blijkt dat er geen significant verschil bestaat tussen de resultaten van 'Taalpeiling yn Fryslân' en het PPON-onderzoek. Zowel de scores van de Friestalige (7.25) als van de Nederlandstalige leerlingen (7.36) van 'Taalpeiling yn Fryslân' blijken niet te verschillen met het PPON-gemiddelde ($F=.848$, $df=2,1006$, $p=.429$). Uit de analyse blijkt tevens dat de scores van de basisscholen

in de onderscheiden categorieën (A, A', B, C en C') niet significant verschillen ($F=.409$, $df=5,1003$, $p=.843$) van de PPON-score op deze luistervaardigheidstoets.

3.1.2 Technische leesvaardigheid

Voor het meten van de technische leesvaardigheid Nederlands is gebruik gemaakt van de Eén-Minuut-Test die voor het PPON-onderzoek is ontwikkeld. Deze lijst bestaat uit 96 woorden. In het verantwoordingsrapport van de PPON-taalpeiling (Zwarts & Rijlaarsdam, 1991, p. 65 e.v.) wordt verantwoord waarom en hoe deze nieuwe toets geconstrueerd is en waarom in het PPON-onderzoek geen gebruik gemaakt is van de bestaande Eén-Minuut-Test van Brus en Voeten (1973). De woordenlijst van de PPON-toets is opgenomen in het 'Voorbeeldenboek' bij 'Balans van het taalonderwijs aan het eind van de basisschool' (PPON, 1991).

Uit de variantie-analyse blijkt dat noch de thuistaal ($F=1.740$, $df=1,196$, $p=.189$) noch de taalomgeving ($F=.938$, $df=2,196$, $p=.393$) voor verschillen zorgen. In een tweede variantie-analyse is nagegaan of het gebruik van het Fries als voertaal en de tijdsinvestering voor het vak Fries samenhangen met de resultaten van de technische leesvaardigheid Nederlands. Voor het aspect gebruik van het Fries als voertaal zijn de scores van de leerlingen van de categorieën A en A' met elkaar vergeleken. Er blijkt sprake te zijn van een significante interactie-effect dat optreedt bij de variabelen thuistaal en voertaalgebruik. Uit de simple-effect analyses blijkt dat de Nederlandstalige leerlingen van categorie A met een gemiddelde van 60.10 significant hoger scoren dan de Friestalige leerlingen van deze categorie, die een score van 52.59 behalen ($F=4.62$, $df=1,124$, $p=.034$). Deze leerlingen blijken bovendien ook significant hoger te scoren dan de Nederlandstalige leerlingen van categorie A', die een gemiddelde score van 48.64 behalen ($F=11.61$, $df=1,124$, $p=.001$).

Voor het aspect tijdsinvestering voor het vak Fries zijn de scores van de leerlingen van de categorieën C en C' met elkaar vergeleken. Hoewel de leerlingen van categorie C ($n=44$) met een gemiddelde van 51.30 woorden per minuut hoger scoren dan de leerlingen van categorie C' ($n=50$) met een gemiddelde van 49.68

woorden per minuut, blijkt uit de variantie-analyse dat dit verschil niet significant is ($F=.362$, $df=1,92$, $p=.549$). Er is dus geen samenhang tussen de tijdsinvestering voor Fries en de technische leesvaardigheid Nederlands.

De gemiddelde PPON-score op technische leesvaardigheid is 53.13. Uit variantie-analyse blijkt dat noch de scores van de Friestalige noch die van de Nederlandstalige leerlingen significant verschillen van de PPON-score ($F=.434$, $df=2,1025$, $p=.648$). Ook de gemiddelde scores van de onderscheiden categorieën basisscholen verschillen niet significant van het PPON-gemiddelde ($F=1.499$, $df=5,1022$, $p=.187$).

3.1.3 Begrijpend lezen

Voor begrijpend lezen is gebruik gemaakt van een deel van de PPON-toetsen met rapporterende teksten (Folder Kinderrechtswinkel en krantebericht over vandaalgevoelige telefooncellen). De toetsen zijn opgenomen in het 'Voorbeeldenboek' bij 'Balans van het taalonderwijs aan het einde van de basisschool' (PPON, 1991).

Uit de analyse blijkt dat er binnen de categorieën A, B en C geen samenhang met de thuistaal van de leerlingen vastgesteld kan worden ($F=2.560$, $df=1,581$, $p=.110$). Ook de verschillen tussen de onderscheiden categorieën basisscholen blijken niet significant te zijn ($F=1.575$, $df=2,581$, $p=.208$).

Om na te gaan of er een samenhang is tussen de resultaten van begrijpend lezen Nederlands enerzijds en het gebruik van het Fries als voertaal en de tijdsinvestering voor het vak Fries anderzijds, zijn de resultaten van de categorie A met die van categorie A' en de resultaten van categorie C met die van categorie C' vergeleken. De resultaten van de variantie-analyse maken duidelijk dat de scores van de leerlingen van categorie A niet significant verschillen van de gemiddelde scores van de leerlingen van categorie A' ($F=.150$, $df=1,286$, $p=.699$). Dit betekent dat er geen samenhang is met de mate van voertaalgebruik Fries. Er wordt wel eens beweerd, dat de resultaten van het begrijpend lezen Nederlands negatief beïnvloed worden door een grotere tijdsinvestering voor het vak Fries. Uit de variantie-analyse blijkt echter dat

het verschil tussen de scores van de Friestalige leerlingen van categorie C', de basisscholen met een dubbele tijdsinvestering voor het vak Fries (7.51), niet significant verschilt van de scores van de leerlingen van de basisscholen van categorie C (7.74) ($F=.883$, $df=1,249$, $p=.348$). Er blijkt dus geen samenhang te zijn tussen de mate van tijdsinvestering voor het vak Fries enerzijds en de resultaten op de toets begrijpend lezen Nederlands anderzijds.

In tegenstelling tot de luistervaardigheid en de technische leesvaardigheid blijkt uit variantie-analyse dat bij begrijpend lezen zowel Friestalige als Nederlandstalige leerlingen uit het onderzoek 'Taalpeiling yn Fryslân' significant hoger scoren dan de leerlingen van het PPON-onderzoek ($F=25.134$, $df=2,1536$, $p=.000$). Uit de analyse blijkt dat alle categorieën, behalve categorie C', eveneens significant hoger scoren dan de leerlingen uit het PPON-onderzoek ($F=9.643$, $df=5,1533$, $p=.000$).

3.1.4 Spreekvaardigheid

Uit eigen waarnemingen in de onderwijspraktijk komt naar voren dat op sommige basisscholen in overwegend Friestalige situaties het produktieve mondelinge gebruik van het Nederlands niet frequent geoefend wordt. In de nota 'Haadlinen foar in taalûnderwiis op de basisskoolen yn Fryslân' (Permanente Kommissje Frysk, 1989, p. 31) wordt aangegeven dat het noodzakelijk is dat de basisscholen in Friesland een duidelijk beleid formuleren voor het voertaalgebruik binnen hun school.

Uit Zwarts (1990, p. 108) blijkt dat op de basisscholen in Nederland als geheel aan gericht spreekonderwijs gemiddeld 24 minuten per week wordt besteed. Deze cijfers beklemtonen de noodzaak van een goed voertaalbeleid binnen de basisscholen in Friesland. Als behalve bij de spreekbeurten bij andere vakgebieden geen consequent gebruik van het Nederlands als voertaal wordt geëist, betekent dit dat Friestalige leerlingen in overwegend Friestalige situaties slechts een half uur per week gericht met het produktieve mondelinge gebruik van het Nederlands bezig zijn. Het is dan ook de vraag of in dergelijke situaties wel voldoende rekening wordt gehouden met de kerndoelen voor het spreken van het Nederlands, zoals deze geformuleerd zijn in het domein 'Spreken

en luisteren' (Besluit kerndoelen 1993).

Om de spreekvaardigheid te toetsen is gebruik gemaakt van de PPON-toets 'Vervelende mensen' (PPON, 1991, p. 114). De leerlingen kregen bij deze taak de oproep van 'Radio Piepjong' voorgelegd om een verhaal te leven over een vervelende gebeurtenis.

De resultaten van de spreekopdracht zijn beoordeeld op een schaal van 0-200. De normtekst was vastgesteld op 100.

De Nederlandstalige leerlingen ($n=52$) van de basisscholen van de categorieën A, B en C halen met een score van 86.56 wel een hoger resultaat dan de Friestalige leerlingen ($n=54$), die 84.11 punten halen. Het verschil is echter niet significant ($F=3.156$, $df=1,100$, $p=.079$). Dit betekent dat er voor de categorieën A, B en C geen sprake is van een samenhang tussen de communicatieve spreekvaardigheid Nederlands en de variabele 'thuis taal'.

De scores van de leerlingen van de onderscheiden categorieën basisscholen liggen dicht bij elkaar. De leerlingen van categorie A ($n=36$) hebben een score van 85.44, die van categorie B ($n=34$) scoren gemiddeld 84.84 en de leerlingen van categorie C ($n=36$) behalen een score van 85.62. De verschillen zijn niet significant ($F=.123$, $df=2,100$, $p=.885$). Dit betekent dat er geen samenhang is tussen de taalomgeving van de basisscholen enerzijds en de communicatieve spreekvaardigheid Nederlands anderzijds. De leerlingen van basisscholen in een overwegend Friestalige situatie scoren op het gebied van de spreekvaardigheid Nederlands dus niet lager dan hun leeftijdgenoten van basisscholen uit een overwegend Nederlandstalige omgeving.

Ten aanzien van de mate van voertaalgebruik geeft de variantie-analyse aan dat er geen significante effecten optreden ($F=.513$, $df=1,75$, $p=.476$).

Uit de variantie-analyse die is uitgevoerd ten aanzien van de tijdsinvestering voor het vak Fries, blijkt dat er geen significante effecten optreden. Dit betekent dat er geen samenhang is tussen de mate van tijdsinvestering voor het vak Fries en het niveau van communicatieve spreekvaardigheid Nederlands.

De gemiddelde PPON-score is 85.36. Als de gemiddelde scores van de Friestalige (84.21) en Nederlandstalige leerlingen (86.60)

met de PPON-score vergeleken worden, dan blijkt dat noch de resultaten van de Friestalige leerlingen noch die van de Nederlandstalige leerlingen significant verschillen van de gemiddelde score van de leerlingen uit het PPON-onderzoek ($F=.403$, $df=2,287$, $p=.669$). Ook de scores van de leerlingen van de onderscheiden categorieën basisscholen zijn vergeleken met de resultaten van het PPON-onderzoek. De uitgevoerde variantie-analyse geeft aan dat geen van de categorieën basisscholen van 'Taalpeiling yn Fryslân' significant verschilt van de gemiddelde score van het PPON-onderzoek ($F=.205$, $df=5,284$, $p=.960$). Dit betekent dat leerlingen van basisscholen in een overwegend Friestalige situatie op het gebied van de spreekvaardigheid Nederlands niet lager scoren dan hun leeftijdgenoten van basisscholen uit een overwegend Nederlandstalige omgeving.

Een nadere beschouwing van de gemiddelde scores laat zien dat zowel het PPON-gemiddelde als de gemiddelde scores van de Friestalige en Nederlandstalige leerlingen in alle categorieën beneden de normscore van 100 liggen. Op grond van deze constatering kan de vraag gesteld worden of er in het basisonderwijs voldoende aandacht is voor de communicatieve spreekvaardigheid zoals omschreven in de kerndoelen Nederlandse taal.

3.1.5 Spellingsvaardigheid

Het door ons gebruikte deel van de toets voor de spellingsvaardigheid Nederlands is in het peilingsonderzoek van PPON door 720 leerlingen gemaakt. De gecorrigeerde score van deze leerlingen is 10.08. In 'Taalpeiling yn Fryslân' is deze toets gemaakt door 783 leerlingen.

Uit de analyse blijkt dat de Friestalige leerlingen in de categorieën A, B en C ($n=338$) een gemiddelde score van 9.97 behalen. De Nederlandstalige leerlingen in deze categorieën ($n=255$) scoren gemiddeld 10.07. Er is dan ook geen sprake van samenhang met de variabele thuis taal ($F=.430$, $df=1,587$, $p=.515$). Uit de analyse blijkt dat ook bij de variabele taalomgeving geen sprake is van een significant effect ($F=.664$, $df=2,587$, $p=.515$). Het voorgaande betekent dat er bij de categorieën A, B en C geen samenhang is tussen de variabelen thuis taal en taalomgeving enerzijds en de resultaten

van de toets spellingsvaardigheid Nederlands anderzijds.

Uit de analyse blijkt eveneens dat er geen enkele samenhang is tussen de spellingsvaardigheid Nederlands en het gebruik van het Fries als voertaal ($F=1.100$, $df=1,287$, $p=.295$). Ook de tijdsinvestering voor het vak Fries blijkt niet samen te hangen met de spellingsvaardigheid Nederlands ($F=.076$, $df=1,251$, $p=.782$).

Uit variantie-analyse van de resultaten blijkt dat de scores van zowel de Friestalige als de Nederlandstalige leerlingen van het onderzoek 'Taalpeiling yn Fryslân' niet significant verschillen van de gemiddelde score van de leerlingen van het PPON-onderzoek ($F=.851$, $df=2,1501$, $p=.427$).

Ook de scores van de onderscheiden categorieën basisscholen zijn vergeleken met de resultaten van het PPON-onderzoek. Uit de analyse blijkt dat geen enkele categorie van basisscholen van het onderzoek 'Taalpeiling yn Fryslân' significant verschilt van het gemiddelde resultaat van de leerlingen die aan het PPON-onderzoek hebben deelgenomen ($F=.707$, $df=5,1498$, $p=.619$).

3.1.6 Communicatieve schrijfvaardigheid

Uit Zwarts (1990, p. 35) blijkt dat volgens de leerkrachten in de basisschool per week ongeveer 25 minuten aan schrijfvaardigheid van zakelijke teksten (stellen) wordt besteed. Dat komt neer op 6% van de totale tijd die per week aan taalonderwijs wordt besteed. Aan het spellingsonderwijs wordt daarentegen per week 20% (83 minuten) van de beschikbare tijd voor taalonderwijs besteed. Kennelijk hechten leerkrachten een grotere waarde aan spellingsvaardigheid dan aan communicatieve schrijfvaardigheid.

Voor het onderdeel communicatieve schrijfvaardigheid is gebruik gemaakt van de toets 'Smikkelreep', opgenomen in het 'Voorbeeldenboek' bij 'Balans' (PPON, 1991, p. 95). De leerlingen moeten in een briefje aan de fabrikant van de smikkelrepen beargumenteren dat zij recht hebben op een reclame-aanbieding.

De resultaten van de schrijfp opdracht zijn beoordeeld op een schaal van 0 - 200. De aangeboden normtekst was vastgesteld op 100.

Hoewel de Nederlandstalige leerlingen van de categorieën A, B en C ($n=96$) met een gemiddelde van 104.24 hoger scoren dan de Friestalige leerlingen ($n=100$) met een gemiddelde van 97.93, is dit verschil niet significant ($F=2.074$, $df=1,190$, $p=.151$). Ook de taalomgeving levert geen significant verschil op ($F=.027$, $df=2,190$, $p=.973$).

Ten aanzien van de mate van gebruik van het Fries als voertaal blijkt uit de variantie-analyse dat de Nederlandstalige leerlingen van categorie A', basisscholen waar het Fries niet als voertaal gebruikt wordt, het laagst scoren met een gemiddelde van 89.47. De Friestalige leerlingen van deze categorie daarentegen scoren met een gemiddelde van 105.00 het hoogst. Uit de analyse blijkt echter dat er geen significante hoofdeffecten en interactie-effecten optreden ($F=.564$, $df=1,117$, $p=.733$).

Ten aanzien van de tijdsinvestering blijken de Friestalige leerlingen van categorie C met een gemiddelde van 100.12 hoger te scoren dan de Friestalige leerlingen van categorie C', die een gemiddelde score van 98.45 behalen. Uit de variantie-analyse blijkt echter dat het verschil niet significant is ($F=.067$, $df=1,91$, $p=.796$). Dit betekent dat er eveneens geen samenhang is met de tijdsinvestering voor het vak Fries.

Uit de variantie-analyse blijkt dat de gemiddelde scores van de Friestalige en de Nederlandstalige leerlingen in Friesland niet significant afwijken van de gemiddelde score van 9.65 van het PPON-onderzoek ($F=.395$, $df=2,476$, $p=.674$).

De scores van de in 'Taalpeiling yn Fryslân' onderscheiden categorieën basisscholen zijn vergeleken met de resultaten van het PPON-onderzoek. Uit de variantie-analyse blijkt dat geen van de categorieën significant verschilt van het landelijke gemiddelde, zoals gemeten in het PPON-onderzoek ($F=.393$, $df=5,473$, $p=.854$).

3.2 De taalvaardigheid Fries

Voor een goed begrip van de resultaten met betrekking tot de taalvaardigheid Fries is het noodzakelijk de volgende gegevens in de beschouwing te betrekken. De tijdsinvestering voor Fries blijft op de meeste basisscholen beperkt tot 45 minuten per week. Een rekenom

leert, dat onderwijs in het vak Fries in de groepen 3 tot en met 8 gedurende 45 minuten per week leidt tot een totale tijdsinvestering van 180 uur gedurende de gehele basisschoolperiode. Dat is nauwelijks 3% van de totale onderwijstijd van de groepen 3 tot en met 8. Een tweede gegeven dat van belang is, betreft het toepassen van differentiatie in doelstellingen. Hoewel het percentage gemengdtaalige scholen sterk is toegenomen, blijkt dat 70% van de basisscholen in Friesland geen differentiatie in doelstellingen tussen Friestalige en niet-Friestalige leerlingen voor het onderwijs in het Fries toepast. Bovendien maken de scholen voor Fries geen onderscheid tussen technisch en begrijpend lezen.

Voor de resultaten van de taalvaardigheid Fries volstaan wij met een overzicht van de taakstelling en het feitelijke niveau van de taalvaardigheid Fries van de leerlingen aan het einde van de basisschool. De taakstelling is gebaseerd op de algemene doelstelling van de kerndoelen Fries dat zowel de Friestalige als de niet-

Friestalige leerlingen het Fries doelmatig moeten kunnen gebruiken in situaties die zich in het dagelijks leven voordoen. In Tabel 1 zijn de taakstelling en de conclusies ten aanzien van het niveau van de beheersing van de deelvaardigheden Fries weergegeven.

Op grond van het beoordelingscriterium dat de leerlingen in dagelijks voorkomende situaties het Fries doelmatig moeten kunnen gebruiken (geconcretiseerd in de door ons gebruikte toetsen), is er sprake van een behoorlijk niveau van beheersing van de luistervaardigheid en het begrijpend lezen door beide groepen leerlingen. Ten aanzien van de communicatieve spreekvaardigheid is alleen bij de Friestalige leerlingen sprake van een behoorlijk niveau. Het niveau van de spreekvaardigheid van Nederlandstalige leerlingen is laag. De deelvaardigheden technisch lezen, spelling en communicatieve schrijfvaardigheid worden door beide groepen leerlingen slechts in geringe mate beheerst. Dit betekent dat er sprake is van een grote discrepantie tussen de geformuleerde

Tabel 1
Taakstelling en feitelijk niveau van de taalvaardigheid Fries

Taakstelling	Friestalige leerlingen	Nederlandstalige leerlingen	Conclusie
<i>Luisteren:</i> Een schoolradio-programma kunnen volgen	88% van de opgaven goed beantwoord	83% van de opgaven goed beantwoord	goed niveau
<i>Technisch lezen:</i> parallel aan deciel 5 groep 8 van de Nederlandse EMT (75-77 woorden) (Brus & Voeten 1973)	58.74 woorden; deciel 1 van groep 8	47.85 woorden; deciel 1 van groep 8	laag niveau
<i>Begrijpend lezen:</i> 70% van de opgaven van de leestekst en de cloze-toets	73% van de leestekst en 76% van de cloze-toets goed beantwoord	69% van zowel de leestekst als de cloze-toets goed beantwoord	behoorlijk niveau
<i>Spreekvaardigheid:</i> normscore van 100 op de spreekvaardigheidstoets	20% van de leerlingen heeft een score van 100 of hoger; 40% heeft een score in de range van 90-100	70% van de leerlingen heeft een score lager dan 90	Nederlandstalige leerlingen een laag niveau; Friestalige leerlingen een behoorlijk niveau
<i>Spelling:</i> beheersing 11 spellingsgevallen van de toets	50% van de fout gespelde woorden wordt herkend	45% van de fout gespelde woorden worden herkend	laag niveau
<i>Schrijfvaardigheid:</i> normscore van 100 op de schrijfvaardigheidstoets	score van 61.0	score van 56.8	laag niveau

kerndoelen voor de Friese taal en de feitelijke taalvaardigheid Fries van zowel Friestalige als Nederlandstalige leerlingen.

Uit de analyses van technisch lezen Fries blijkt dat met name bij Nederlandstalige leerlingen het gebruik van het Fries als voertaal een positief effect heeft op de technische leesvaardigheid Fries. Nederlandstalige leerlingen op basisscholen waar het Fries niet als voertaal gebruikt wordt, scoren significant lager dan Nederlandstalige leerlingen op basisscholen in een vergelijkbare taalsituatie waar het Fries wel als voertaal een plaats heeft. Dit heeft ongetwijfeld te maken met het feit dat door het gebruik van het Fries als voertaal de woordenschat van de leerlingen vergroot wordt, waardoor een snellere woordherkenning mogelijk gemaakt wordt. Dit gegeven komt overeen met een onderzoek naar de resultaten van het onderwijs in het Baskisch (Sierra & Olaziregi, 1991). Uit dat onderzoek blijkt dat het onderwijsmodel waarin het Baskisch alleen als vak wordt onderwezen geen hoge resultaten voor de taalbeheersing Baskisch oplevert. Met name het gebruik van de tweede taal als voertaal zorgt voor een betere ontwikkeling van de vaardigheid in de tweede taal.

4 Slotbeschouwing

Uit het onderzoek 'Taalpeiling yn Fryslân' blijkt, dat zowel de Friestalige als de Nederlandstalige leerlingen in Friesland aan het einde van de basisschool hetzelfde niveau van taalbeheersing in het Nederlands bereiken als de leerlingen op de basisscholen in Nederland als geheel. Deze conclusie is een bevestiging van conclusies uit eerder en elders verricht onderzoek naar de beheersing van het Nederlands van Friestalige kinderen (Wijnstra, 1976) resp. van dialectsprekende kinderen (Kuijper, Stijnen & Van den Hoogen, 1983) en van Turkse kinderen in Nederland (Teunissen, 1986). Deze conclusie vindt bevestiging in buitenlandse literatuur over de samenhang tussen onderwijs in de minderheidstaal en de beheersing van de meerderheidstaal: "Data from bilingual programmes all over the world show that instruction through the minority language does not hinder children's academic development in

the majority language" (Cummins, 1989, p. 21).

Uit ons onderzoek blijkt, dat de resultaten van de taalbeheersing Fries minder positief te noemen zijn. Hoewel de kerndoelen voor Fries volledig parallel geformuleerd zijn aan de kerndoelen voor Nederlands, blijkt uit het onderzoek 'Taalpeiling yn Fryslân', dat noch de Friestalige leerlingen noch de Nederlandstalige leerlingen op de basisscholen in Friesland deze kerndoelen aan het einde van de basisschool in voldoende mate bereiken.

De verklaring voor dit verschil in niveau van taalbeheersing tussen Nederlands en Fries is te vinden in de in het schoolwerkplan geformuleerde doelstellingen en in de gebruikelijke tijdsinvestering voor Fries. Verreweg de meeste basisscholen in Friesland besteden niet meer dan één klokuur per week aan Fries als vak, waarbij de aandacht vooral uitgaat naar lezen. Door deze aanpak kunnen de leerlingen aan het einde van de basisschool uiteraard niet in alle vier de taalvaardigheden in beide schooltalen hetzelfde niveau bereiken. Vastgesteld moet worden, dat het tweetalig onderwijs in Friesland nog niet gericht is op het bereiken van volledige tweetaligheid.

Om te komen tot verantwoord tweetalig onderwijs in Friesland is het noodzakelijk een raamplan voor een geïntegreerde benadering van het onderwijs in het Nederlands en in het Fries te ontwikkelen. Hierbij kan gebruik gemaakt worden van het reeds ontwikkelde raamplan voor een geïntegreerde methode voor NT1- en NT2-onderwijs (Sijstra, 1994). In bedoeld raamplan moet in elk geval aandacht besteed worden aan de volgende aspecten:

- "(heterogene) beginsituatie op school- en klassenniveau
- toepassing van modellen voor tweetalig onderwijs
- transfer Nederlands-Fries: samenhang moedertaalonderwijs en memmetaalonderwijs
- woordenschat Nederlands en Fries
- beleid t.a.v. allochtone leerlingen

Onder invloed van demografische ontwikkelingen in de laatste decennia (migratie en 'gemengde' huwelijken) zijn er in Friesland vrijwel geen homogeen Friestalige gebieden of

plaatsen meer, maar ook geen homogeen niet-Friestalige gebieden of plaatsen. Als gevolg van deze ontwikkelingen kennen alle basisscholen in meer of mindere mate een gemengd-talig samengestelde schoolbevolking, variërend van minder dan 20% Friestalige leerlingen tot meer dan 80% Friestalige leerlingen. Uit De Jong (1986) blijkt echter, dat scholen tot nu toe nauwelijks differentiëren op grond van de thuistaal van de leerlingen. In het raamplan moeten derhalve verschillende modellen van tweetalig onderwijs uitgewerkt worden voor de verschillende taalsituaties. Gemeenschappelijk kenmerk van deze modellen is, dat het vak Fries en het vak Nederlands in dezelfde klas zowel op het niveau van moedertaalonderwijs als op het niveau van tweede-taalonderwijs functioneren. De samenstelling van de schoolbevolking naar thuistaal is hierbij richtinggevend voor de mate waarin Fries en Nederlands als moedertaal resp. als tweede taal onderwezen worden. Dit vraagt om een geïntegreerde benadering van het taalonderwijs, gericht op het verwerven van alle taalvaardigheden in beide talen. Hierbij kan en moet nadrukkelijk gebruik gemaakt worden van didactische mogelijkheden gericht op transfer tussen beide talen. Basisvaardigheden kunnen in beginsel in één van beide talen aangeleerd worden en vervolgens hun toepassing vinden in de andere taal. Bovendien moet de didactiek erop gericht zijn, dat niet-taalgebonden vaardigheden die in de ene taal geleerd zijn, tevens toegepast worden in de andere taal.

De uitbreiding van de woordenschat vormt een belangrijk aspect van de taalontwikkeling van jonge kinderen (zie ook Leseman & Şahin, dit nummer). Een uitgebreide en adequate woordenschat is een noodzakelijke voorwaarde voor schoolsucces en voor het functioneren in de tweetalige maatschappij. In het basisonderwijs zou aandacht besteed moeten worden aan het systematisch verwerven van een basiswoordenschat in beide talen.

Ten slotte is het van belang expliciet aandacht te besteden aan het onderwijs aan allochtone leerlingen. Ook in Friesland is sprake van een toename van het aantal allochtone leerlingen, die in de komende jaren steeds verder verspreid over de provincie deel zullen nemen aan het onderwijs. Tot nu toe is noch in het regionale on-

derwijsbeleid noch in de didactiek noch in de materiaalontwikkeling rekening gehouden met de aanwezigheid van allochtone leerlingen. Het is voor een evenwichtige ontwikkeling van het taalonderwijs in Friesland van groot belang, dat op beleidsniveau en op schoolniveau een antwoord geformuleerd wordt op de vraag of de verplichting onderwijs in het Fries te geven ook voor allochtone leerlingen moet gelden en zo ja, aan welke voorwaarden dit onderwijs dan zou moeten voldoen.

Noten

- 1 Na het afsluiten van het manuscript van dit artikel is een nieuw survey gepubliceerd in 'Taal yn Fryslân op 'e nij besjoen' (Gorter & Jonkman, 1995). De actuele cijfers over de taalbeheersing Fries luiden: 94% van de inwoners kan Fries verstaan, 74% kan Fries spreken, 65% kan Fries lezen en 17% kan Fries schrijven.
- 2 Deze gegroeide praktijk is in 1995 gecodificeerd in de Algemene Wet Bestuursrecht (hoofdstuk 2) en in de Wet Gebruik Friese taal in het Rechtsverkeer.

Literatuur

- Besluit kerndoelen (1993). *Besluit van 4 mei 1993, houdende vaststelling van de kerndoelen basisonderwijs (Besluit kerndoelen basisonderwijs)*. Staatsblad 264.
- Brus, B. Th., & Voeten, M. J. M. (1973). *Eén-Minut-Test. Verantwoording en handleiding*. Nijmegen: Berkhout.
- Cummins, J. (1989). Language and literacy acquisition in bilingual contexts. In *Journal of Multilingual and Multicultural Development*, 10, 17-31.
- Gorter, D., Jelsma, G. H., Plank, P. H. van der, & Vos, K. de (1984). *Taal yn Fryslân. Undersyk nei taalgedrach en taalhâlding yn Fryslân*. Ljouwert: Fryske Akademy.
- Inspectie van het onderwijs (1989). *Het onderwijs in het Fries op de basisschool: stand van zaken 1988-1989. Een kwantitatief inspectieonderzoek naar het onderwijs in het Fries op de basisscholen in Friesland*. Zoetermeer.
- Jong, S. de (1986). *Het onderwijs in het Fries nader bekeken*. Inspectierapport nr. 2. Den Haag: Staatsdrukkerij.

- Jong, S. de, & Riemersma, A. M. J. (1994). *Taalpeiling yn Fryslân. Onderzoek naar de beheersing van het Fries en het Nederlands aan het einde van de basisschool*. Academisch proefschrift KUB, Tilburg. Ljouwert: Fryske Akademy.
- Kuijper, H., Stijnen, S., & Hoogen, J. van den (1983). *Onderwijs tussen dialect en standaardtaal*. Nijmegen: NCDN.
- Ministerie van C.R.M. (1970). *Rapport van de Commissie Friese-taalpolitiek*. Den Haag.
- Permanente Kommissje Frysk (1989). *Haadlinen foar it taalûnderwiis op de basisskoolen yn Fryslân*. Ljouwert.
- PPON (1991). *Balans van het taalonderwijs aan het einde van de basisschool. Voorbeeldenboek*. Arnhem: CITO.
- Sierra, J., & Olaziregi, I. (1991). *E.I.F.E.3. Influence of factors on the learning of Basque. Study of the models A, B and D in second year Basic General Education*. Gasteiz, Spain: Central Publications Service of the Basque Government.
- Sijtsma, J. M. (1994). *Taal voor iedereen. Raamplan voor een geïntegreerde NT1/NT2-methode voor de basisschool*. 's Hertogenbosch: K.P.C.
- Teunissen, J. M. F. B. G. (1986). *Eén school, twee talen. Een onderzoek naar de effecten van een tweetalig-bicultureel onderwijsprogramma voor Marokkaanse en Turkse leerlingen op basisscholen in Enschede*. Academisch proefschrift. Utrecht: RUU.
- Wijnstra, J. M. (1976). *Het onderwijs aan van huis uit Friestalige kinderen. Verslag van een evaluatie-onderzoek in een meertalige regio*. Academisch proefschrift. Den Haag: SVO.
- Zwarts, M. (1990). *Balans van het taalonderwijs aan het einde van de basisschool*. PPON-reeks nr. 2. Arnhem: CITO.
- Zwarts, M., & Rijlaarsdam, G. (red.) (1991). *Verantwoording van de taalpeiling einde basisonderwijs 1988*. Arnhem: Cito.

Auteurs

S. de Jong is inspecteur van het basis- en (voortgezet) speciaal onderwijs in Friesland.

Adres: Rijksinspectiekantoor (RIK), Postbus 933, 8901 BS Ljouwert/Leeuwarden.

A. M. J. Riemersma is hoofd van de Halbertsma-Akademie (lerarenopleidingen Fries van de NHL en de CHN te Leeuwarden) en stafmedewerker van het 'Berie foar it Frysk', adviesorgaan van Gedeputeerde Staten van Friesland inzake het Friese-taalbeleid.

Adres: Berie foar it Frysk, Postbus 20120, 8900 HM Ljouwert/Leeuwarden.

Abstract

Language Assessment in Friesland

S. de Jong & A. M. J. Riemersma, Pedagogische Studiën, 1996, 73, 291-302.

The study 'Language Assessment in Friesland' ('Taalpeiling yn Fryslân') aims to establish the level of fluency attained in both of the languages used in schools in Friesland. The formal organization of the study is largely based on the model employed in the national survey 'Periodical Assessment of Educational Levels' (PPON). By consequence of this approach, a comparison of the results of both research projects has become possible. The analysis of the results shows that there is no connection between the level of the pupil's fluency in Dutch and either the language spoken in the pupil's home or the language environment of the school he or she attends. The pupil's results with regard to oral and reading comprehension in Frisian can be considered satisfactory. Their performance in the areas of basic (technical) reading and spelling was insufficient and remained behind the desired level with regard to speaking and writing.

Studievaardigheden in de brugklas. Een vergelijking tussen algemene en vakspecifieke verwerkingsstijlen van brugklasleerlingen*

E. B. Klatter**

Samenvatting

In dit artikel wordt een onderzoek beschreven naar studievaardigheden van brugklasleerlingen. Nagegaan is welke relatie bestaat tussen enerzijds leeractiviteiten die leerlingen uitvoeren in een concrete vakopdracht en anderzijds testgegevens die deze activiteiten navragen. Twee instrumenten zijn daarvoor gebruikt: de Inventaris Leerstijlen (Vermunt, 1992) en een Vakspecifieke Inventaris Leerstijlen Test, binnen het vak Nederlands-tekstverklaring. Duizend brugklasleerlingen hebben beide tests gemaakt. Factor- en correlatie-analyses laten zien dat er weinig samenhang bestaat tussen de gegevens van de conceptueel gelijke tests. Vervolgens zijn relaties bepaald tussen de testgegevens en rapportcijfers voor het vak Nederlands. De verwerkingsactiviteiten die leerlingen zeggen uit te voeren blijken niet samen te hangen met hun schoolse prestaties, terwijl de vakspecifieke opdracht met name significante verbanden laat zien met het rapportcijfer. Deze resultaten hebben geleid tot een nadere studie van de Inventaris Leerstijlen-items. Vraaggesprekken met twintig brugklasleerlingen over een aantal items expliciteerden onduidelijkheden in de test. Deze onduidelijkheden kunnen onder andere oorzaak zijn van de eerder verkregen resultaten en zouden tevens een indicatie kunnen vormen voor een inhoudelijke aanpassing van het instrument.

Inleiding

De laatste tien jaar is door onderwijspsychologisch onderzoek duidelijk aangetoond dat verschillen in leerprestaties tussen leerlingen voor een belangrijk deel kunnen worden verklaard door de wijze waarop kennis is verworven, en door de mate waarin leerlingen in staat zijn hun eigen leerprocessen te sturen (Wang, Haertel & Walberg, 1990). Mede dankzij deze inzichten is de aandacht voor studievaardigheidstrainingen sterk toegenomen. Daarbij wordt getracht adequate leerstrategieën aan te leren. In dit verband wordt ook wel gesproken over 'leren leren'. Om tot volwaardig leren te komen, zijn o.a. cognitieve en metacognitieve activiteiten vereist. De eerste categorie activiteiten betreft de concrete leeractiviteit zelf, zoals het memoriseren van woordjes, of het maken van een samenvatting van een stuk tekst. Metacognitieve activiteiten betreffen de sturingsactiviteiten van de concrete leerhandeling. Bijvoorbeeld, de keuze om een les sec uit het hoofd te leren, of om toch eerst een samenvatting te maken.

Metacognitieve activiteiten krijgen binnen het studievaardigheids-onderwijs relatief veel aandacht. Flavell (1979) omschrijft metacognitie als de kennis en de controle die iemand heeft over de eigen cognitieve processen. Elshout (1990) spreekt over dit soort kennis als processen van een 'hogere orde', daarmee aangevend dat lagere-orde-processen slechts adequaat kunnen verlopen als de hogere-orde-processen ook goed verlopen. Deze omschrijvingen verwijzen naar de overstijgende eigenschap van metacognitie, waarbij uitspraken als 'hogere-orde-processen' en 'controlerende functies' het belang ervan illustreren voor het leerproces.

Toch kunnen vraagtekens gesteld worden bij de grotere aandacht voor metacognitieve ac-

* Dit onderzoek is uitgevoerd aan de Rijksuniversiteit Groningen. Een eerdere versie van het kwantitatieve deel is gepubliceerd in de proceedings van de Landelijke Dag Studievaardigheden '95.

** Met dank aan G. Tillemans en G. G. H. Jansen voor de begeleiding. In het bijzonder dank aan J. D. H. M. Vermunt voor de zinvolle suggesties en collegiale medewerking.

tiviteiten ten opzichte van de cognitieve activiteiten. Een vraag bijvoorbeeld is in hoeverre de cognitieve activiteiten kunnen worden gereguleerd als geen zicht bestaat op de al of niet aanwezigheid ervan. Een leerling kan bijvoorbeeld plannen om een schema te maken van een stuk tekst. Als vervolgens blijkt dat hij of zij geen goede schema's kan maken, zal de voorafgaande, goede metacognitieve keuze niet per definitie tot een goede cognitieve activiteit en vervolgens tot een goed leerproduct leiden.

Om een beeld te krijgen van de cognitieve activiteiten waarover leerlingen beschikken, is getracht deze activiteiten op vakspecifiek niveau in kaart te brengen.

Vermunt (1992) heeft een instrument ontwikkeld, de Inventaris Leerstijlen (ILS), waarmee o.a. het gebruik van cognitieve en metacognitieve activiteiten in kaart kan worden gebracht. De ILS is een diagnostisch instrument waarmee een *algemene leerstijl* wordt gemeten. (Dit instrument en het begrip leerstijl worden verderop in dit artikel beschreven.) Om de uitgevoerde cognitieve activiteiten op vakspecifiek niveau in kaart te brengen, is naast de ILS een Vakspecifieke Inventaris Leerstijlen Test (VILT) ontworpen. Deze situatie bood de mogelijkheid het onderzoek een validerend karakter mee te geven. De VILT is daarom op een dusdanige wijze geconstrueerd dat ze conceptueel identiek is aan de ILS.

Tegen de achtergrond van bovenstaande gedachten presenteert dit artikel de uitkomsten van een onderzoek naar de cognitieve activiteiten van brugklasleerlingen. Voor het onderzoek zijn de volgende doelen geformuleerd:

- * het ontwikkelen van een instrument aan de hand waarvan cognitieve activiteiten op vakinhoudelijk domein (Nederlands) in kaart kunnen worden gebracht van leerlingen uit de brugklas van het voortgezet onderwijs,
- * nagaan of de cognitieve activiteiten die met de ILS als componenten van een algemene leerstijl kunnen worden onderscheiden, terug te vinden zijn in de manier waarop leerlingen een vakspecifieke opdracht uitvoeren.

Ten slotte krijgt de kwantitatieve benadering, bij wijze van derde onderzoeksdoel, aanvulling in een aantal kwalitatieve interviews. Naast de validiteitsvraag is achteraf de vraag gesteld naar de betrouwbaarheid van de ILS bij de po-

pulatie van leerlingen uit de brugklas, aangezien de gebruikte ILS-versie voor brugklasleerlingen (Van Rijswijk, Vermunt, De Jong & Kluvers, 1991) een variant is van de oorspronkelijke vragenlijst voor studenten.

1 Theoretische achtergrond

In deze paragraaf wordt nader ingegaan op het theoretische begrippenkader van leerstijlen. In het toonaangevende werk van Vermunt (1992) op dit gebied, wordt op heldere wijze een theoretische onderbouwing gegeven. Het door hem ontwikkelde instrument om leerstijlen te meten, de ILS, alsmede de systematische indeling van componenten van leerstijlen, zijn daarom als fundament gekozen voor de verdere uitwerking van het onderzoek.

In paragraaf 1.1 wordt ingegaan op de inhoud van de begrippen cognitieve en metacognitieve activiteiten. In paragraaf 1.2 wordt het begrip leerstijl verder uitgewerkt.

1.1 Leeractiviteiten

In het kader van *studievaardigheden* worden leeractiviteiten de laatste jaren vaak besproken. Volgens Vermunt (1992) zijn leeractiviteiten onder te brengen in drie hoofdcategorieën: de cognitieve verwerkingsactiviteiten, de metacognitieve regulatie activiteiten en de affectieve leeractiviteiten. Omdat de laatste categorie, de affectieve activiteiten, buiten het onderzoek blijft, wordt hier niet verder op ingegaan. De hoofdcategorieën van leeractiviteiten worden elk weer onderverdeeld in een achttal subcategorieën. (1) De *cognitieve verwerkingsactiviteiten* omvatten relateren, structureren, analyseren, concretiseren, toepassen, memoriseren en herhalen, kritisch verwerken en selecteren. Tabel 1 geeft een nadere omschrijving van deze leeractiviteiten. Het zijn alle activiteiten die direct zijn gericht op het verwerken van leerstof en informatie. Zij beogen een directe toeëigening van elementen uit de leerstof zoals feiten, concepten, procedures en principes. Het uitvoeren van deze functies leidt tot leerresultaten in de vorm van begrip, kennis, vaardigheden of inzicht of fysiek constateerbare producten als aantekeningen, samenvattingen en dergelijke. (2) De *metacognitieve leeractiviteiten* sturen en controleren de

Tabel 1**Cognitieve leeractiviteiten (Uit: Vermunt, 1992)****Relateren**

Zoeken naar verbanden tussen verschillende onderdelen van de leerinhoud, tussen de onderdelen en het geheel, de grote lijn van de studiestof en tussen nieuwe informatie en de eigen voorkennis of preconcepties.

Structureren

Samenbrengen van afzonderlijke stukken informatie in een georganiseerd geheel, proberen structuur aan te brengen in de lesstof en integreren van nieuw verworven kennis waarover men reeds beschikt.

Analyseren

Opsplitsen van een groter geheel in de onderdelen waaruit het is samengesteld, stap voor stap uitzoeken welke verschillende aspecten aan een probleem, gedachtengang of theorie te onderkennen zijn.

Concretiseren

Zich concrete voorstellingen proberen te maken bij abstracte informatie, ontleend aan verschijnselen die al bekend zijn.

Toepassen

Zich oefenen in het gebruiken van leerinhouden.

Memoriseren en herhalen

Inprenten van afzonderlijke stukken informatie door deze een aantal malen voor zichzelf te repeteren, en herhalen van de studiestof.

Kritisch verwerken

Meedenken met auteurs, docenten en medestudenten, een eigen mening hebben en niet zomaar alles accepteren van wat geschreven staat of wordt gezegd.

Selecteren

Onderscheiden van hoofd- en bijzaken, reduceren van grote hoeveelheden informatie tot de belangrijkste onderdelen.

cognitieve (en affectieve) leeractiviteiten. Er is sprake van controle en sturing van de eigen leeractiviteiten wanneer de leerling deze activiteiten flexibel en in de juiste situatie kan uitvoeren. Vermunt onderscheidt acht metacognitieve leeractiviteiten: oriënteren, plannen, proces bewaken, toetsen, diagnostiseren, bijsturen, evalueren en reflecteren.

1.2 Leerstijl

Een *leerstijl* kan worden gedefinieerd als de combinatie van cognitieve en metacognitieve activiteiten, samen met de studie-opvattingen en studie-motieven. De onderlinge samenhang van deze vier categorieën bepaalt de leerstijl. De begrippen studie-opvattingen en -motieven verwijzen naar de algemene instelling waarmee een student/leerling aan het leerproces deelneemt. Voor nadere uitwerking van deze begrippen verwijs ik naar de dissertatie van Vermunt (1992). De ILS is oorspronkelijk ontworpen voor een populatie van studenten uit het hoger onderwijs om variaties in *leerstijlen* vast te stellen op vakoverstijgend niveau. In het

onderzoek bij studenten konden vier leerstijlen worden onderscheiden: een ongerichte leerstijl, een reproductieve leerstijl, een betekenisgerichte leerstijl en een toepassingsgerichte leerstijl. De laatste twee staan bekend als de meer adequate leerstijlen, waarmee over het algemeen betere studieresultaten worden geboekt. Bij de bespreking van de onderzoeksresultaten in paragraaf 3, zal worden ingegaan op verschillende kenmerken van de afzonderlijke leerstijlen.

2 Methode

Voor het antwoord op de eerste twee onderzoeksvragen zijn twee instrumenten afgenomen. De ILS-brugklasversie is gebruikt, waarbij het in eerste instantie ging om de gegevens van één schaal: de algemene cognitieve activiteiten. Om deze activiteiten bij een vakspecifieke opdracht na te gaan is de Vakspecifieke Inventaris Leerstijlen Test (VILT) afgenomen. Voor beantwoording van de derde onderzoeksvraag

vraag zijn interviews gehouden met brugklasleerlingen met betrekking tot diverse ILS-items. Dit deel wordt in paragraaf 4 beschreven, na bespreking van de resultaten van het kwantitatieve onderzoeksdeel.

2.1 Instrumenten

De *ILS-brugklasversie (ILS-bk)* is een leerdiagnostisch instrument (Van Rijswijk, Vermunt, De Jong & Kluvers, 1991) en is een aangepaste ILS-versie voor brugklasleerlingen als variant op de oorspronkelijke ILS voor studenten. Met dit instrument worden naast de cognitieve en metacognitieve activiteiten (deel A) ook de studie-opvattingen en studie-motieven (deel B) in kaart gebracht. De ILS is opgebouwd uit de delen A en B, en bevat in totaal 120 items. De items bestaan uit stellingen over (A) manieren van, en (B) opvattingen over leren. Op een vijfpunts-schaal kan een leerling zijn mening geven over de mate waarin een bepaalde stelling op hem van toepassing is (A), of het al dan niet eens te zijn met een uitspraak (B). Voor dit onderzoek is de totale ILS afgenomen, maar bij de analyses is voornamelijk gebruik gemaakt van dat deel van A dat betrekking heeft op de cognitieve activiteiten. Deze activiteiten, in totaal 27 items, kunnen door gedeeltelijke samenvoeging worden gereduceerd tot vijf schalen: relateren/structureren (7 items), kritisch verwerken (4 items), memoriseren (5 items), analyseren (6 items), concretiseren/toepassen (5 items). De items kunnen daarnaast worden ondergebracht in drie subschalen op basis waarvan het uiteindelijke (leerstijl)profiel voor deze activiteiten wordt vastgesteld: (1) diepteverwerking: relateren, structureren en kritisch verwerken; (2) stapsgewijze verwerking: analyseren en memoriseren; en (3) concrete verwerking: concretiseren en toepassen.

De VILT is een domein-specifieke opdracht binnen het vak Nederlands, waarbij om twee redenen is gekozen voor *tekstverklaring*. Ten eerste mag worden aangenomen dat alle leerlingen ervaring hebben met het bestuderen van teksten, en ten tweede biedt deze vorm de mogelijkheid om de acht cognitieve activiteiten na te vragen. De VILT bestaat uit een tekst met een vragenlijst van in totaal 34 vragen. De tekst is ontleend aan de Clibtoetsen van het CITO,

met als onderwerp 'het oog', beslaat anderhalve pagina tekst (A4-formaat, regelafstand 1,5) en staat qua moeilijkheidsgraad volgens de Cito-normen met een Clib-score van 63 op een gemiddeld leesniveau van groep 8 van het basisonderwijs. De vragenlijst bestaat uit open vragen en wordt in de testsituatie pas gemaakt ná bestudering en inlevering van de tekst, inclusief eventueel gemaakte aantekeningen. Goed beantwoorde vragen worden met een punt gehonoreerd, terwijl een nul wordt gescoord voor foutieve of niet beantwoorde vragen. Dat betekent dat een maximale score van 34 punten kan worden behaald.

Bij het opstellen van de vragen hebben de cognitieve activiteiten zoals die in Tabel 1 zijn omschreven, als uitgangspunt gefunctioneerd. De vragen zijn opgesteld in overleg met Vermunt, om een zo groot mogelijke consensus te bereiken over de interpretatie van de omschreven activiteiten. Het beantwoorden van de vragen veronderstelt het uitvoeren van een bepaalde cognitieve leeractiviteit. Opgemerkt moet worden dat het hier om een *veronderstelling* gaat; het is immers nooit precies vast te stellen welke mentale handeling in het hoofd van een lerende plaatsvindt. Door middel van bepaalde vraagstellingen, afgezet tegen de tekstinhoud, is getracht bepaalde cognitieve activiteiten uit te lokken. Hieronder volgen enkele voorbeelden van 'corresponderende' vragen uit beide tests.

Een ILS-vraag betreffende de activiteit 'relateren' ziet er als volgt uit: *Ik probeer te bedenken wat de onderwerpen uit verschillende lessen van een vak (i.c. delen uit een tekst) met elkaar te maken hebben.* Deze activiteit is voor de VILT geoperationaliseerd aan de hand van de volgende vraag: *Wat heeft een lichtbron, uit de eerste alinea, te maken met de kegeltjes uit de laatste alinea?* Om deze vraag te kunnen beantwoorden moet een leerling een relatie leggen tussen de werking van kegeltjes en een lichtbron (kegeltjes functioneren alleen goed onder invloed van een lichtbron waardoor kleuren kunnen worden waargenomen). Een ILS-vraag met betrekking tot 'toepassen' luidt als volgt: *Als ik buiten school een probleem tegenkom, probeer ik dat op te lossen met wat ik op school geleerd heb.* Een equivalent daarvan voor de VILT luidt: *Wat is het gevolg als iemand maar*

twee typen kegeltjes heeft? Het antwoord vereist het toepassen van de kennis over het aantal typen kegeltjes en de functie ervan (elk type is gevoelig voor één kleur, de combinatie van de drie typen zorgt voor alle kleurschakeringen, dus: gedeeltelijk kleurenblind). Een voorbeeld van een VILT-vraag met betrekking tot de activiteit 'analyseren': *Welk onderdeel van ons oog richt het licht dat in een oog valt, op de receptoren?* Om het antwoord op deze vraag te kunnen geven (niet letterlijk in de tekst) moet een leerling 'het oog' stap voor stap langs gaan en analyseren welke weg of welk proces een lichtbundel 'doorloopt'. Een algemene ILS-vraag omtrent 'analyseren' luidt: *Als de leraar uitlegt hoe je een bepaalde som moet maken pluis ik alle stappen die je moet uitvoeren één voor één uit.* Met betrekking tot de 'memoriserende' activiteit is naar feiten gevraagd, zoals de volgende VILT-vraag: *Wat is een fovea en waar bestaat het uit?* De antwoorden staan letterlijk in de tekst (gele vlek, bestaat alleen uit kegeltjes). Een ILS-vraag op dit gebied is bijvoorbeeld: *Rijtjes met kenmerken van bepaalde onderwerpen leer ik uit mijn hoofd.*

Voor de cognitieve activiteiten relateren, structureren, toepassen, memoriseren, kritisch werken en selecteren zijn ieder vier vragen geformuleerd en voor analyseren en concretiseren elk vijf. De vragen zijn derhalve onder te brengen in de drie subschalen diepteverwerking, stapsgewijze verwerking en concrete verwerking (vgl. ILS).

De aanpak, die met de VILT gegeven is levert mogelijk op voorhand een vertekening op. Gezien het onderwerp 'tekstverklaring' lijken goede lezers een voorsprong te krijgen boven zwakke lezers. Uit onderzoek van Voeten, Bergmans, Aarnoutse en De Groot (1992) naar 'tekststructuur en reproductie van informatieve teksten' blijkt dat goede lezers niet alleen meer structuurbesef hebben dan zwakke lezers, maar ook beter in staat zijn om de informatie uit de tekst te reproduceren. Zowel ten aanzien van hoofdideeën als ten aanzien van details vertonen de goede lezers een hogere mate van reproductie. Voeten e.a. (1992) vinden evenwel tegen hun verwachting in geen significante interactie tussen het leesniveau en het niveau van tekstbegrip. Dit gegeven rechtvaardigt het

gebruik van de voor de VILT gekozen aanpak.

Ten aanzien van voorkennis van de leerlingen, is (achteraf) gevraagd of zij bekend waren met de materie van de tekst. Dit bleek voor geen enkele groep het geval te zijn. Veel leerlingen gaven aan de tekst goed te begrijpen, maar de vragen 'wel moeilijk' te vinden.

2.2 Proefpersonen

Voor de samenstelling van de steekproef zijn diverse scholen benaderd met of zonder een gedifferentieerd brugjaar. Leerlingen van 5 LBO/MAVO brugklassen, 4 MAVO brugklassen en 31 MAVO/HAVO/VWO brugklassen hebben deelgenomen. Omdat de aanpassing van de ILS voor brugklasleerlingen heeft plaatsgevonden op brede scholengemeenschappen, vormden de verschillende onderwijsniveaus in principe geen belemmering voor dit onderzoek. Er participeerden acht scholen, met in totaal 1048 leerlingen. Echter, niet alle leerlingen hebben beide tests gemaakt; 984 leerlingen hebben de ILS ingevuld, terwijl 994 leerlingen de VILT hebben gemaakt.

2.3 Procedure

De ILS en de VILT werden allebei in hun geheel afgenomen; de afname van de ILS nam gemiddeld een half uur in beslag, de VILT gemiddeld één lesuur. Tijdens de mondelinge instructie van de VILT werd de leerlingen nadrukkelijk verteld dat aan de uitkomsten van de test geen consequenties verbonden zouden worden; noch voor het rapportcijfer, noch voor de schoolkeuze.

De beschikbare tijd was maximaal één lesuur. Met betrekking tot het leren van de tekst en het maken van de vragenlijst waren de leerlingen vrij in de verdeling van de tijd. Tijdens het leren van de tekst kon een woordenboek worden gebruikt en was het mogelijk om aantekeningen te maken. Zowel de tekst als de aantekeningen moesten bij aanvang van de vragenlijst ingeleverd worden.

2.4 Dataverwerking

Voor het verwerken van de data zijn de volgende analyses uitgevoerd: betrouwbaarheidsanalyse, factor-analyse en correlatiematrix-analyse. De betrouwbaarheid is bepaald met behulp van Cronbachs Alpha (α). Profielen van leerstijlen zijn zowel voor de ILS als voor de

VILT vastgesteld met een geforceerde vari-max-oplossing (hoofcomponenten-analyse) in vier factoren. Om de samenhang tussen de leerstijlprofielen van de ILS en de VILT na te gaan, zijn correlaties berekend tussen de scores op de sub- en hoofdschalen van beide instrumenten. Daarnaast zijn correlaties berekend tussen enerzijds de rapportcijfers voor het vak Nederlands en anderzijds de ILS- en VILT-scores.

3 Resultaten van het kwantitatieve onderzoek

De ILS-brugklasversie is afgenomen bij 984 leerlingen. De betrouwbaarheden van de vijf schalen zijn over het geheel genomen redelijk tot goed: de coëfficiënten variëren van .51 tot .77. Daarmee zijn de uitkomsten nagenoeg identiek aan de resultaten van het onderzoek waarin de ILS werd aangepast voor de brugklas (Kluvers & Simons, 1992). De VILT is gemaakt door 994 leerlingen. Hoewel de betrouwbaarheidscoëfficiënten van de afzonderlijke schalen van dit instrument een stuk lager uitvallen dan die van de ILS-bk, α varieert van .25 tot .55, is de betrouwbaarheid van de VILT, met een α van .78, in zijn geheel genomen redelijk goed. De moeilijkheidsgraad van de totale test, uitgedrukt in de p -waarde is .37; indien de laatste 8 items worden weggelaten, die veelal door tijdgebrek niet zijn ingevuld en derhalve een vertekening opleveren, dan stijgt p tot .41 en daalt de moeilijkheidsgraad. De p -waarden van de drie hoofdschalen diepte, stapsgewijze en concrete verwerking (zonder de 8 items) zijn respectievelijk .36, .44 en .42. Hieruit blijkt dat vragen die beroep doen op een diepere verwerkingsvorm moeilijker worden gevonden dan vragen die een oppervlakkige verwerking veronderstellen. Dit bleek eveneens uit de evaluerende gesprekjes, waarin leerlingen kenbaar maakten dat "die vragen niet in de tekst stonden, daar moest je zelf over nadenken en dat is moeilijk."

De testcores van de ILS-bk met betrekking tot de cognitieve verwerkingsactiviteiten worden bij een drie-factoren-analyse verdeeld over de drie subschalen diepte, stapsgewijze en concrete verwerkingsactiviteiten. Bij een vier-factor-model met betrekking tot de totale ILS-

scores, komen de factoren in grote lijnen overeen met de leerstijlen zoals die door Vermunt zijn onderscheiden bij studenten uit het hoger onderwijs. De resultaten van deze Varimax-analyse zijn weergegeven in Tabel 2.

Tabel 2
Factorladingen van ILS-schalen in een 4-factoren Varimax oplossing ($N = 984$; ladingen $> .25$ en $< .25$ zijn weggelaten)

ILS-schalen	F1	F2	F3	F4
Verwerkingsstrategieën				
Diepte verwerking				
- relateren/structureren	.82	.27		
- kritisch verwerken	.75			
Stapsgewijze verwerking				
- memoriseren/herhalen	.25	.79		
- analyseren	.39	.76		
Concrete verwerking	.68			
Regulatiestrategieën				
Zelfsturing				
- leerproces en -res.	.60	.48		
- leerinhoud	.67	.25		
Externe sturing				
- leerproces		.52		
- leerresultaten		.46	.40	
Stuurloos leergedrag				.75
Studie-opvattingen				
Opbouw van kennis	.53	.35	.44	
Opname van kennis		.45	.57	.34
Gebruik van kennis	.53		.58	
Stimulerend onderwijs	.27		.48	.50
Samen studeren				.62
Studiemotieven				
Persoonlijke interesse	.38		.58	
Certificaatgericht		.34	.65	
Testgericht	.31		.61	
Beroepsgericht			.71	
Ambivalent				.75
Eigenwaarde	8.21	2.01	1.64	1.46
Perc. verklaarde variantie	37.3	9.1	7.4	6.7
Idem cumulatief	37.3	46.4	53.9	60.6

De eerste factor representeert de *betekenisgerichte leerstijl*. Leerlingen bij wie deze leerstijl overheerst maken uit zichzelf gebruik van diepte verwerkingsactiviteiten zoals relateren, structureren en kritisch verwerken. Daarnaast geven zij aan hun leerprocessen, -resultaten en -inhouden zelf te sturen. Ze zien leren als het opbouwen en gebruiken van kennis, waarbij zij vooral door persoonlijke interesse in een vakgebied aangemoedigd worden, hetgeen de studie-opvattingen en -motieven vertegenwoordigen. Leerlingen bij wie een *reproductie-gerichte leerstijl* – de tweede factor –

domineert, gebruiken tijdens het leren met name activiteiten als memoriseren, herhalen en analyseren. Hun leerprocessen en -resultaten worden voornamelijk extern gestuurd, terwijl hier evenals bij de betekenisgerichte leerstijl een deel van de sturing door de leerling zelf wordt uitgevoerd. Studenten met een reproductieve leerstijl zien leren vooral als het opnemen van kennis. Opmerkelijk is daarom dat voor brugklasleerlingen niet alleen het opnemen van kennis, maar ook de opbouw van kennis binnen deze dimensie is vertegenwoordigd. De derde factor kan eigenlijk *niet* worden geïnterpreteerd als een *toepassingsgerichte leerstijl*. Voor deze leerstijl is bij studenten een hoge lading gevonden voor de concrete verwerkingsactiviteit (Vermunt, 1992). Bij brugklasleerlingen echter vinden we geen lading voor deze verwerkingsactiviteit. Ook voor de sturing van de leerresultaten worden brugklasleerlingen extern aangedreven in tegenstelling tot studenten. Zij vatten leren op in de meest brede betekenis van het woord; zowel het opbouwen, het opnemen en het gebruik van kennis, alsmede het stimulerend onderwijs zijn in hun studieopvattingen vertegenwoordigd. In hun studiemotieven zijn deze leerlingen gericht op het halen van certificaten en het testen van hun eigen mogelijkheden. Daarnaast zijn zij persoonlijk geïnteresseerd en overeenkomstig het profiel van studenten, overheerst bij leerlingen binnen deze leerstijl de beroepsgerichte oriëntatie.

Uit deze gegevens kan geconcludeerd worden dat het leergedrag van brugklassers kennelijk nog maar in beperkte mate wordt gestuurd door de manier waarop zij over leren denken en wat zij ermee willen bereiken. Ten slotte kan de vierde factor worden getypeerd als een *ongegerichte leerstijl*. Leerlingen met overwegend deze leerstijl maken geen of nauwelijks gebruik van leeractiviteiten en regulatie van het leergedrag blijft geheel achterwege. Sturingsactiviteiten worden niet door henzelf geïnitieerd en zij maken geen gebruik van sturing die uitgaat van de instructie. Uit de ladingen voor studie-opvattingen blijkt dat deze leerlingen vinden en verwachten dat het onderwijs erop gericht moet zijn om hen te stimuleren. Ook het samen studeren wordt hoog aangeslagen. Conform de verwachting hebben deze leerlingen een ambivalente leeroriëntatie (Roosendaal, 1993).

Niet alle kenmerken van leerstijlen, zoals die zijn gevonden bij studenten uit het hoger onderwijs, zijn terug te vinden in het overzicht van Tabel 2. Het meest opvallende is dat de toepassingsgerichte leerstijl in feite nog niet als afzonderlijke leerstijl voorkomt, maar min of meer ondergedompeld is in de betekenisgerichte stijl. Tevens blijkt bij brugklasleerlingen nog weinig differentiatie aanwezig in de studie-opvattingen en studie-motieven (Vermunt, 1995). Vergelijken we deze categorieën met de resultaten van de onderzoeken van Roosendaal (1993) en Vermunt (1992), met respectievelijk VWO-5 leerlingen en WO-studenten, dan valt op dat brugklasleerlingen over de verschillende leerstijlen heen, leren opvatten als een mengeling van het opbouwen van kennis, kennis opnemen en het gebruiken van kennis. Praktisch alle leerlingen, verdeeld over de vier leerstijlen zien het (in meer of mindere mate) als taak van het onderwijs om hen te stimuleren tot het uitvoeren van leeractiviteiten. Deze gegevens zijn niet geheel verrassend als naast de 'zich ontwikkelende bewustwording' van deze leerlingen, het verplichte karakter van het voortgezet onderwijs in oogenschouw wordt genomen. Immers, leerlingen kiezen niet zelfstandig of bewust voor dit onderwijs, waardoor weinig kans bestaat dat reeds bepaalde opvattingen of motieven om aan dat onderwijs deel te nemen, uitgekristalliseerd zijn.

Factor-analyses van de *VILT-scores* leverden minder eenduidige resultaten op. Bij een acht factoren factor-analyse werden de items zonder samenhang over de acht factoren verspreid. Een factor-analyse met drie factoren, waarbij de items naar verwachting over de subschalen diepte, stapsgewijze en concrete verwerking verdeeld zouden moeten worden, leverde eveneens weinig structuur op. Opvallend was dat bij deze analyse de derde factor uitsluitend werd gevormd door de laatste negen items van de vragenlijst. Bij nadere bestudering van de data bleek dat veel leerlingen nul scoorden op de laatste 8 à 10 vragen. Deze vragen zijn veelal niet meer gemaakt in verband met de tijd. Doordat deze items logischerwijs met elkaar correleren, levert dat een factor op die kan worden geïnterpreteerd als 'tijdgebrek'. Een drie-factoren-model waarbij de scores op de laatste tien vragen werden weggelaten, leverde

een verschuiving van de resterende items op. De derde factor werd nu gevormd door (de laatste) vier items van de eerste factor uit de oorspronkelijke analyse. De items representeren weer verschillende cognitieve activiteiten. Ook nu kon de factoroplossing niet binnen de theoretische uitgangspunten worden geïnterpreteerd.

Mogelijke samenhang tussen de scores op de ILS en de VILT is daarom nagegaan op grond van de theoretische indeling in subschalen en hoofdschalen. De opzet van de VILT is dusdanig, dat de items van de vragenlijst zowel kunnen worden geordend naar de acht cognitieve verwerkingsactiviteiten, als naar de vijf schalen van leeractiviteiten die in de ILS voorkomen. Als de ILS en de VILT hetzelfde meten, zouden de uitkomsten van een schaal van de ILS, bijvoorbeeld analyseren, sterk moeten samenhangen met die VILT-vragen die deze schaal vertegenwoordigen. Om inzicht te verkrijgen in deze samenhang zijn de correlatiecoëfficiënten berekend tussen zowel de opgetelde itemscores van de subschalen als van de hoofdschalen van de VILT en de ILS.

Tabel 3
Correlaties tussen subschalen en hoofdschalen van de ILS en de VILT

Subschaal	correlatie ILS - VILT	N
relateren/structureren	.072	848
kritisch verwerken	.019	900
analyseren	.049	883
memoriseren	-.033	906
concretiseren	.024	906
Hoofdschaal		
diepteverwerking	.067	826
stapsgewijze verwerking	-.073	861
concrete verwerking	.024	906

De coëfficiënten zijn, zoals uit Tabel 3 valt af te lezen, uitzonderlijk laag en het verband is non-existent. Nadere exploratie leverde nog wel enige samenhangen op. De totaalscores van de VILT werden gerelateerd aan totaalscores van de drie overige ILS-categorieën; de sturingsactiviteiten, studiemotieven en -opvattingen. Verbanden zijn te constateren tussen de totaalscores van de VILT en twee subcategorieën van de ILS-bk, te weten: 'stuurloos' binnen de sturingsactiviteiten en 'ambivalent' binnen de studiemotieven, met correlatie-coëffi-

ciënten van respectievelijk $-.25^{**}$ en $-.17^{**}$ ($N=731$, significantieniveau: $^{**} < .001$ tweezijdige toetsing).

De samenhangen tussen de VILT-scores en de twee subcategorieën van de ILS-bk zijn zwak maar significant. Logischerwijs zijn beide coëfficiënten negatief: naarmate een leerling hoger scoort op de VILT, vertoont hij of zij minder stuurloze en ambivalente leeractiviteiten en andersom.

Voor deze teleurstellende resultaten ligt de verklaring voor de hand dat de VILT een ondeugdelijk instrument is en niet die activiteiten toetst, die ze zou moeten toetsen. Om te bezien of deze verklaring opgaat, is het predictieve vermogen van de test nagegaan, door als externe factor de studieresultaten in de analyse op te nemen. Immers, als cognitieve leeractiviteiten geacht kunnen worden ergens een bijdrage aan te leveren, dan toch tenminste aan de studieresultaten. Voor dat doel zijn van 490 leerlingen uit de steekproef, afkomstig van vier scholen, de rapportcijfers voor het vak Nederlands verzameld. De correlatiecoëfficiënten in Tabel 4 laten de verbanden zien tussen het laatst behaalde rapportcijfer en de scores op de ILS- en VILT-schalen.

Er bestaan geen significante samenhangen tussen het laatst behaalde rapportcijfer voor het vak Nederlands en de cognitieve verwerkingsactiviteiten van de overigens intern consistente ILS-bk. De cognitieve activiteiten van de minder eenduidige VILT, vertonen enkel positief significante verbanden met het rapportcijfer. Ook de subschalen van de VILT leveren interpreteerbare verbanden op. De uitkomst met betrekking tot de VILT ligt eigenlijk voor de hand; deze testvorm komt meer overeen met de gangbare toetsvorm voor het vak Nederlands, dan een instrument als de ILS. Ook het feit dat de subschaal memoriseren hoog correleert met het cijfer voor Nederlands ($r = .35^{**}$) ligt in de lijn der verwachting. In de brugklas geldt over het algemeen dat naarmate een leerling veel memoriserende en herhalende verwerkingsactiviteiten uitvoert, hij of zij in staat is betere cijfers te behalen (Palinscar & Winn, 1990). Het verschil tussen goede en zwakke lezers komt in het al genoemde onderzoek van Voeten e.a. (1992) sterker tot uiting in de reproductie van details dan in de reproductie van hoofdideeën. Het lijkt erop dat in het brugjaar niet zozeer de

Tabel 4

Correlaties tussen ILS-schalen/VILT-schalen met rapportcijfer voor Nederlands ($N = 490$, ladingen $> -.05$ en $< .05$ zijn weggelaten)

	Rapportcijfers	
	ILS	VILT
Verwerkingsstrategieën		
Diepte verwerking		.23**
- relateren/structureren		.17**
- kritisch verwerken		.17**
Stapsgewijze verwerking		.32**
- memoriseren/herhalen		.35**
- analyseren	-.11	.18**
Concrete verwerking	-.10	.26**
Regulatiestrategieën		
Zelfsturing		
- leerproces en -res.	-.07	
- leerinhoud	-.12	
Externe sturing		
- leerproces	-.13*	
- leerresultaten	-.09	
Stuurloos leergedrag	-.38**	
Studie-opvattingen		
Opbouw van kennis		
Opname van kennis	-.06	
Gebruik van kennis	-.05	
Stimulerend onderwijs	-.08	
Samen studeren		
Studiemotieven		
Persoonlijk geïnteresseerd	.05	
Certificaatgericht	-.08	
Testgericht	-.08	
Beroepsgericht	.06	
Ambivalent	-.38**	

significantiëniveau: * -.01 ** -.001 (tweezijdige toetsing)

diepere verwerking van informatie een voorwaarde is voor schoolsucces, als wel het goed kunnen memoriseren en onthouden.

Wat betreft de metacognitieve activiteiten die met de ILS zijn gemeten, is er een negatief significante samenhang te constateren tussen het rapportcijfer en de externe sturing van het leerproces. Daarnaast is een relatief sterke, negatieve samenhang gevonden tussen het rapportcijfer en de mate van stuurloos leergedrag. Het is duidelijk dat deze relatie negatief is; een leerling zal beter presteren naarmate hij of zij beter weet op welke wijze een leertaak het best kan worden aangepakt. Voor de categorie studie-opvattingen bestaat geen enkel significant verband met schoolse prestaties voor Nederlands. De cijfers voor dit vak vertonen wel een sterke samenhang met ambivalente studie-

motieven. Logischerwijs is ook deze samenhang negatief.

De totaalscores van de VILT vertonen een positief significante relatie met prestaties voor het vak Nederlands ($r = .35^{**}$, $N=490$). Deze gegevens impliceren dat de VILT een redelijke representant is van schoolse leerprestaties. De mate waarin leerlingen binnen de ILS-bk aan-geven bepaalde activiteiten uit te voeren, vertoont daarentegen weinig verband met het rapportcijfer. De eventuele ondeugdelijkheid van de VILT, geïndiceerd door de zwakke factor-structuur, lijkt daarmee geen afdoende verklaring te bieden voor de ontbrekende samenhang tussen ILS- en VILT-scores.

4 Opzet en resultaten van het kwalitatieve gedeelte

De lage correlaties tussen de ILS en de VILT, en de afwezigheid van enige samenhang tussen rapportcijfers en de ILS-scores van de cognitieve activiteiten, vormden aanleiding om de ILS op itemniveau nauwkeuriger te bekijken. Twee mogelijke oorzaken van de ontbrekende verbanden tussen de ILS en de VILT en tussen de ILS en de rapportcijfers zijn in stellingen uitgewerkt. Deze stellingen functioneerden als uitgangspunt voor de derde onderzoeksvraag en tevens het kwalitatieve deel van het onderzoek (een aantal kwalitatieve interviews met brugklasleerlingen).

De eerste stelling luidde: "Het is niet denkbeeldig dat een 12/13-jarige leerling minder reflectief gedrag vertoont dan een volwassene." Juist bij het invullen van een vragenlijst als de ILS wordt beroep gedaan op kennis over het eigen leergedrag. Dit soort kennis, ook wel metacognitieve kennis, ontstaat veelal door reflectie op het eigen leergedrag (Elshout-Mohr & Van den Bijtel, 1995). De veronderstelling is dat leerlingen van deze leeftijd *niet spontaan* reflecteren op de activiteiten die zij ondernemen. Als daardoor niet precies duidelijk is welke activiteiten een leerling (per taak) uitvoert, neemt de kans toe dat een vragenlijst als de ILS redelijk 'willekeurig' wordt ingevuld. Het ontbreken van *metacognitieve kennis* kan op die manier debet zijn aan het feit dat een 'vertekende leerstijl' wordt verkregen.

De tweede stelling zoekt het op een lager ni-

veau in de begrijpelijkheid en eenduidigheid van de vraagstelling: "Zowel mondelinge als schriftelijke communicatie valt of staat bij een juiste interpretatie van de boodschap." Het invullen van de ILS kan worden gezien als een vorm van schriftelijke communicatie, waarbij leerlingen gevraagd wordt uitspraken omtrent leren en onderwijs te beantwoorden. De *interpretatie* van deze uitspraken is cruciaal voor het adequaat beantwoorden van de lijst. Verschillende concepten van de participerende partijen (leerlingen én testontwerpers) kunnen leiden tot verschillende interpretaties ten aanzien van zowel woorden, uitspraken of zinsconstructie als ten aanzien van de uit te voeren activiteiten.

Vervolgens zijn interview-sessies georganiseerd met het doel een beeld te krijgen van de mate waarin leerlingen over metacognitieve kennis beschikken omtrent hun eigen leeractiviteiten. Daarnaast is getracht om eventuele verschillen in interpretaties van de participanten met betrekking tot de ILS-items in kaart te brengen.

Op basis van proefinterviews is, in samenwerking met Tijmens (KUN), een definitieve keuze gemaakt van de meest problematische vragen. In totaal zijn 22 items geselecteerd van een verkorte versie van de ILS (69 items), waarna een standaard-interview is ontworpen. Dit interview is afgenomen bij 22 leerlingen uit diverse brugklassen (VBO, MAVO, HAVO en VWO). Tijdens het interview werd elk item door de interviewer hardop voorgelezen, waarna de vraag werd gesteld: "Wat wordt volgens jou met deze vraag bedoeld?". Op deze manier werd de persoonlijke interpretatie van elk item door de leerling geëxpliciteerd. Vervolgens werd doorgevraagd over de specifieke leeractiviteit "Doe jij dat wel eens"; "Kun je een voorbeeld geven"; "Kan dit bij alle vakken of denk je dat je deze aanpak juist bij één bepaald vak gebruikt"; enzovoorts. De interviews zijn op band opgenomen en uitgeschreven. De protocollen zijn geanalyseerd en de antwoorden op individuele items gescoord in drie categorieën: (1) niet of fout begrepen en/of verkeerd geïnterpreteerd; (2) enigszins begrepen maar bij doorvragen blijft het vaag; (3) goed begrepen en geïnterpreteerd. Van het totaal aantal antwoorden op de vragen (100%), week 16% dusdanig af van de veronderstelde intentie

dat een 1 werd gescoord. Voor 12% van de antwoorden werd een 2 gescoord en uit het merendeel van de antwoorden (72%) bleek dat de leerlingen de items goed hadden begrepen en op een juiste manier interpreteerden. In totaal leidde dus 28% van de besproken ILS-items tot dubbelzinnige antwoorden of onbegrip. Sommige begrippen, zoals 'voorbereiding' of 'buiten school' waren voor brugklasleerlingen *multi-interpretabel*. Andere begrippen zoals 'samenbrengen tot een geheel' waren voor hen te moeilijk en begrippen als 'een totaalbeeld' en 'kritisch', bleken geheel onbekend. Juist door een verkeerde interpretatie verandert de inhoud van de vraag. In principe kan alsnog een goed antwoord worden gegeven, maar dit antwoord vertegenwoordigt vervolgens een ander leergedrag dan verondersteld is met de vraag. In dit verschijnsel ligt mogelijk enige grond voor de scheve profielen van de ILS en de ontbrekende verbanden van de ILS met de rapportcijfers en derhalve met de VILT.

De eerste stelling die hierboven is beschreven, betrof het reflectieve gedrag en tevens de metacognitieve kennis van brugklasleerlingen. Deze aspecten werden niet sec onderzocht, maar via de interview-gegevens kan hier het een en ander over worden gezegd. Een aantal leerlingen heeft de 'interview-items' eerst zelfstandig beantwoord, voordat naar de betekenis werd gevraagd. Hierbij viel op dat de leerlingen alle items vlot konden invullen, terwijl er bij doorvraag toch onduidelijkheden bleken te bestaan over 'wat ze eigenlijk echt doen'. De onderliggende gedachte is dat leerlingen *intuïtief* wel een idee hebben van de leeractiviteiten die zij toepassen en prefereren, hierop wijzen de uitkomsten van de ILS, maar deze niet goed kunnen expliciteren en concretiseren. Dit kan een verwijzing zijn naar het achterwege blijven van reflectie, waardoor opbouw van meta- cognitieve kennis weinig kans krijgt. Dit idee wordt tevens ondersteund door het feit dat leerlingen moeizaam voorbeelden konden geven. Naarmate de 'vertaling' van het begrip (de activiteit) beter was, was een voorbeeld sneller gevonden. De voorbeelden die werden gegeven, bleven evenwel zeer concreet en vakspecifiek. Daarnaast was het verwoorden in *kindertaal* voor de leerlingen heel gewoon. De vragen van de ILS worden min of

meer opgevat als 'volwassentaal', hoewel bij de aanpassing van de ILS getracht is de items juist in terminologie van 12/13-jarigen te vertalen. Dit lijkt een 'bevestiging' van de tweede stelling, waarbij de kwantitatieve gegevens van het interview tevens aanduiden dat hier regelmatig sprake is van een vorm van misinterpretatie. Uiteindelijk zeiden veel leerlingen het *moeilijk* te vinden om op deze manier over leren te praten; "we zijn het niet gewend en denken er ook nooit zo over na."

Op grond van deze kwalitatieve gegevens kan de globale conclusie worden geformuleerd, dat leerlingen op het niveau van de brugklas in het algemeen nog weinig gewend zijn boven de concrete leersituatie uit te stijgen en weinig beroep (kunnen) doen op overkoepelende metacognitieve kennis ten aanzien van hun eigen leergedrag. Als zij al hun leergedrag overdenken, blijft dit beperkt tot *domeinspecifieke taken* waarover in concrete termen wordt nagedacht.

5 Discussie

Naar aanleiding van de ILS-resultaten bij een populatie van studenten en brugklasleerlingen, kan worden gesteld dat beide categorieën in grote lijnen dezelfde leerstijlprofielen vertonen. Er zijn echter een paar duidelijke verschillen te constateren. Bij leerlingen van twaalf tot dertien jaar komen niet alle cognitieve verwerkingsactiviteiten naar voren die bij bepaalde leerstijlen passen. Ook is uit de onderzoeksresultaten af te lezen dat brugklasleerlingen diverse vormen van sturing naast elkaar gebruiken, en vergeleken met studenten zijn de studie-opvattingen en studie-motieven van brugklasleerlingen nog heel diffuus.

De profielen van de VILT hebben vergeleken met de profielen van de ILS-bk, een zeer afwijkende vorm. De verbanden tussen de subschalen van de VILT en de ILS zijn uiterst zwak, dan wel afwezig. De lage correlaties kunnen deels beïnvloed zijn door het feit dat de items van de ILS per schaal vrij eenduidig en gelijkvormig zijn, terwijl de vragen van de VILT per schaal zeer divers zijn. Daar staat tegenover dat de VILT-scores een redelijk predictief vermogen hebben ten aanzien van leerresultaten. Een en ander combinerend, ligt de

gedachte voor de hand, dat de items met betrekking tot de cognitieve activiteiten van de ILS-bk en de VILT beroep doen op verschillende soorten activiteiten of vaardigheden. Een conclusie luidt dan ook dat de ILS en de VILT niet representatief zijn voor elkaar; zij meten verschillende hypothetische constructen (vgl. Boekaerts & Simons, 1993).

Die constructen zijn verschillend van aard; de ILS (deel A) doet direct navraag naar het *leerproces*, terwijl met de VILT getracht wordt de aard van cognitieve leeractiviteiten indirect af te leiden uit het *leerproduct*. Ter illustratie het volgende voorbeeld: een leerling geeft in de ILS aan een bepaalde activiteit bijvoorbeeld relateren, veel uit te voeren (ILS-score 5). Hij doet dat in werkelijkheid ook (zoekt naar verbanden tussen alinea's), maar komt daarbij niet tot een herkenbaar resultaat (fout antwoord) met als gevolg een VILT-score van 0. Dit aspect heeft wellicht een rol gespeeld bij de lage overeenkomsten tussen ILS- en VILT-scores. De twee instrumenten brengen in feite verschillende meetproducten in kaart.

Daarnaast duiden de resultaten van de correlatie-analyses tussen de rapportcijfers voor het vak Nederlands en de VILT-scores, op een lineair verband tussen enerzijds de VILT en anderzijds proefwerken en overhoringen binnen de schoolcontext. Het feit dat de ILS-bk ook activiteiten binnen dergelijke schoolse taken navraagt, maar daarmee geen significante relatie vertoont, doet vermoeden dat bij de beantwoording van de ILS-items nog een ander fenomeen meespeelt.

Tijdens de interviewsessies gaven leerlingen aan dat het antwoord op veel vragen kon variëren, al naar gelang welk vak in gedachte werd gehouden. Gedurende de interviews switchten de leerlingen regelmatig van het ene vak naar het andere. Deze *vakspecificiteit* kan sterke invloed hebben op het profiel van de leerstijl. Bij het ene vak kunnen bijvoorbeeld bepaalde activiteiten overheersen, terwijl deze bij een ander vak totaal niet aan bod hoeven te komen (Thomas & Rohwer, 1987; Boekaerts & Simons, 1993). De wisselende aandacht voor leerstofdomeinen heeft waarschijnlijk niet alleen geleid tot onlogische en zelfs tegenstrijdige profielen bij de ILS, maar kan ook de lage correlaties tussen de ILS en een vakspecifieke opdracht als de VILT tot gevolg hebben gehad.

Impliciet en expliciet bieden deze gegevens het antwoord op de onderzoeksvragen. Het is niet gelukt gelijkvormige cognitieve verwerkingsstijlen te destilleren op zowel algemeen als vakspecifiek niveau. De mate waarin eventueel overeenkomsten bestaan tussen algemene en vakspecifieke verwerkingsstijlen is daarmee niet eenduidig vast te stellen. Maar juist het ontbreken van de verbanden leidt evenwel tot nieuwe inzichten.

De vraag of 12/13-jarige leerlingen over voldoende reflectieve vermogens, en daarmee metacognitieve kennis beschikken om de ILS-items 'waarheidsgetrouw' in te kunnen vullen, lijkt ontkennend te kunnen worden beantwoord. Uit de interviews werd duidelijk dat leerlingen in deze leeftijdscategorie niet of nauwelijks gewend zijn expliciet over hun leergedrag te praten. Dit zijn aanwijzingen voor het feit dat binnen de onderwijscontext het bespreken van leergedrag, in tegenstelling tot het leerprodukt nog steeds niet gebruikelijk is. Binnen het studievvaardigheidsonderwijs worden wellicht handvatten geboden om een adequate manier van leren te hanteren, maar de eigen strategieën worden veelal niet eerst geëxpliciteerd en/of gediagnostiseerd. De vraag is dan aan welke kennis leerlingen hun antwoorden ontleen bij het invullen van een vragenlijst over ondernomen leeractiviteiten. De kans is daardoor groot dat er enigszins 'willekeurig' wordt beantwoord en dat zodoende de gevonden leerstijl niet representatief is voor de daadwerkelijke leerstijl.

Deze redenering leidt makkelijk tot de opmerking dat een vragenlijst als de ILS niet adequaat is voor brugklasleerlingen, wegens onvoldoende metacognitieve kennis bij deze doelgroep, maar wel bruikbaar is bij een populatie van studenten die wel over deze kennis zou beschikken. Toch kan dit niet zo simpel worden gesteld. Het probleem is dat zelfrapportage over eigen cognitief functioneren wordt beïnvloed door hetzelfde cognitieve functioneren dat men pretendeert te meten (Riemersma, 1993). Beantwoording van de ILS impliceert nog niet dat ook duidelijk wordt wat studenten in concrete situaties werkelijk doen. Juist bij onderwijspsychologisch onderzoek is dit een belangrijk issue.

Daarnaast is het de vraag of items van de ILS-brugklasversie 'identiek' moeten zijn aan

die van de oorspronkelijke ILS. Want zijn leeractiviteiten die ervaren leerders ondernemen op eenzelfde wijze vertegenwoordigd in het scala van activiteiten van beginnende leerders? De interviews laten zien dat dit niet het geval is. Zo wordt 'woordjes leren', een memoriserende activiteit, veel genoemd door brugklasleerlingen, terwijl deze activiteit niet wordt nagevraagd in de ILS-bk. Ook valt op dat bepaalde vragen niet echt relevant zijn voor het onderwijs in de brugklas. Te denken valt aan de gefaseerde opbouw van methoden, waarbij leerlingen de samenhang van onderwerpen min of meer aangeboden krijgen. Ter illustratie geven we de antwoorden van twee brugklassers op de ILS-stelling: *Ik probeer te bedenken wat de onderwerpen uit verschillende lessen van een vak met elkaar te maken hebben*. De ene leerling constateert: "... we gaan meestal van onderwerp naar onderwerp, dus dan hoort alles wel bij elkaar (...), dan heeft het wel wat met elkaar te maken. Daar ga ik nooit over nadenken." De andere antwoordt met de opmerking: "Nee, doe ik niet, want meestal heb je in het boek wel een bepaalde volgorde dat je snapt waar het over gaat." Uit de antwoorden van deze, overigens slimme leerlingen blijkt dat zij, gegeven de voorstructurering van het onderwijs in de brugklas, op dit terrein weinig relaterende verwerkingsactiviteiten verrichten. Vermeld moet worden dat deze leerlingen zich echter goed bewust zijn van datgene dat zij (juist niet) denken. Zij vormden evenwel de minderheid.

6 Tot slot

Het is gebleken dat de ILS goed functioneert bij een populatie die al geruime tijd onderwijs heeft gevolgd en kennis heeft van het eigen leergedrag binnen diverse onderwijscontexten. Het onderzoek van Vermunt (1992) laat zien dat er samenhangen bestaan tussen leeractiviteiten en persoons- en contextgebonden factoren die een verklaring vormen voor de stabiliteit van leerstijlen. Dit geldt niet voor leerlingen met beduidend minder onderwijservaring, bij wie de studie-opvattingen en -motieven nauwelijks blijken samen te hangen met de gekozen leeractiviteiten. Daarnaast wordt twijfel omtrent de reële weergave van de ondernomen cognitieve leeractiviteiten door

brugklassers in dit onderzoek ondersteund door persoonlijke gesprekken. Deze resultaten leidden tot de gedachte dat de ILS-bk meer een indicatie vormt voor de aan- of afwezigheid van metacognitieve kennis ten aanzien van het eigen leergedrag, dan dat het instrument een algemene leerstijl in kaart brengt.

De onderzoeksvraag naar het bestaan van een relatie tussen algemene en vakspecifieke verwerkingsstijlen kan vervolgens ontkennend worden beantwoord. De gegevens van de ILS-bk vallen niet samen met de leeractiviteiten die een brugklasleerling onderneemt tijdens een vakspecifieke opdracht. Dit blijkt uit vergelijking van de resultaten van beide tests. Vanwege de twijfelachtige aard van de VILT kan niet worden hard gemaakt dat de ILS een meer of minder zuivere weergave van de werkelijkheid is dan de vakspecifieke test. Daarentegen vormen de resultaten van het kwalitatieve onderdeel, bijvoorbeeld de multi-interpretable formulering van de items, toch een indicatie om voorzichtig om te springen met de resultaten van een diagnostisch instrument als de ILS-bk.

Deze algemene conclusie kan consequenties hebben voor de praktijk van geïntegreerde studievoordigheidstraining. Een logische stap zou zijn om tijdens training in studievoordigheid vooral aandacht te besteden aan reflectie. Door te reflecteren kan de leerling zich meer bewust worden van onderliggende cognitieve activiteiten die hij/zij uitvoert. Door inzicht in het eigen leergedrag ontstaat begrip van verschillende interventies. Ons inziens vormen deze aspecten de basis om leerlingen op te leiden tot zelfstandige leerders. Om het 'leren leren' wezenlijk gestalte te geven zou daarom niet alleen aandacht besteed moeten worden aan regulering maar vooral ook aan diagnostisering van de cognitieve activiteiten als bouwstenen van het leerproces.

Literatuur

Boekaerts, M., & Simons, P. R. J. (1993). *Leren en instructie - Psychologie van de leerling en het leerproces*. Assen: Dekker & Van de Vegt.

Elshout, J. J. (1990). Hogere-orde vaardigheden: discussie. In M. J. Ippel & J. J. Elshout (Eds.), *Training van hogere orde denkprocessen* (pp. 149-155). Amsterdam/Lisse: Swets & Zeitlinger.

Elshout-Mohr, M., & Bijtel, J. A. L. van den (1995). Reflecteren: een nuttige ambachtelijke vaardigheid. In H. C. Schouwenburg & J. T. Groenwoud (Red.), *Studievoordigheid en leerstijlen* (pp. 229-246). Groningen: Wolters-Noordhoff.

Flavell, J. H. (1979). Metacognition and cognitive monitoring, a new area of cognitive developmental inquiry. *American Psychologist*, 34, 906-911.

Klatte, E. B. (1995). Leerstijlen in de brugklas - onderzoek naar een vakspecifieke leerstijl. In H. C. Schouwenburg & J. T. Groenwoud (Red.), *Studievoordigheid en leerstijlen* (pp. 169-192). Groningen: Wolters-Noordhoff.

Kluvers, C., & Simons, P. R. J. (1992). *Zelfregulatievaardigheden en COO*. Onderzoeksverslag Vakgroep Onderwijskunde: Katholieke Universiteit Nijmegen.

Palinscar, A. S., & Winn, J. (1990). Introduction to the special issue on assessment models focussed on new conceptions of achievement and reasoning. *International Journal of Educational Research*, 14, 411-413.

Riemersma, F. S. J. (1993). Leerstijlen en procesgerichte instructie. *De Psycholoog*, 28, 367-368.

Roosendaal, L. A. (1993). *Wisselwerking tussen leer- en doceerstrategieën*. Interne publikatie Sectie Onderwijs- en Opleidingspsychologie, Katholieke Universiteit Brabant, Tilburg.

Rijswijk, F. van, Vermunt, J. D. H. M., Jong, F. P. C. M. de, & Kluvers, C. (1991). *Inventaris Leerstijlen voor het voortgezet onderwijs*. Tilburg/Nijmegen: Katholieke Universiteit Brabant/Katholieke Universiteit Nijmegen.

Thomas, J. W., & Rowher, W. D. (1987). Grade-level and course-specific differences in academic studying: summary. *Contemporary Educational Psychology*, 12, 381-385.

Vermunt, J. D. H. M. (1992). *Leerstijlen en sturen van leerprocessen in het hoger onderwijs - Naar procesgerichte instructie in zelfstandig denken*. Lisse: Swets & Zeitlinger.

Vermunt, J. D. H. M. (1995). Leerstijlen: een overzicht en recente onderzoeksgegevens. In H. C. Schouwenburg & J. T. Groenwoud (Red.), *Studievoordigheid en leerstijlen* (pp. 51-72). Groningen: Wolters-Noordhoff.

Voeten, M. J. M., Bergmans, A., Aarnoutse, C. A. J., & Groot, I. de (1992). Effect van tekststructuur op de reproductie van informatieve teksten bij brugklasleerlingen. *Tijdschrift voor Onderwijsresearch*, 17, 238-255.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*, 84, 30-43.

Manuscript aanvaard 9-1-1996

Auteur

E. B. Klatter is Assistent in Opleiding aan de Katholieke Universiteit Nijmegen Vakgroep Onderwijskunde en doet onderzoek naar de (ontwikkeling van) leerconcepties van leerlingen van 12 tot 14 jaar.

Adres: Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen, Postbus 9104, 6500 HE Nijmegen.

Abstract

Domain-specific learning activities related to general learning styles.

E. B. Klatter. *Pedagogische Studiën*, 1996, 73, 303-316.

This study addresses the relationship between general learning activities and domain-specific learning activities as categories of a learning style. The theoretical basis of this research is mainly found in the work of Vermunt (1992). He developed the questionnaire 'Inventory LearningStyles' (ILS) to display general learning styles. The main purpose of this study is to examine the external validity of the ILS. In particular, it is investigated whether the scale 'cognitive learning activities' of the ILS represents the cognitive activities twelve-year-old students really use while studying. For this reason another instrument was developed at a domain-specific level (VILT) with scales representing different cognitive activities. More than one thousand first-graders filled out both tests. Data analyses showed very differing factors for each test and hardly any correlations between the scales of the two instruments. Interviews with 22 first-grade students illustrated mis-interpretations and different conceptions about some ILS-items which might have been reflected in the results.

Boekbespreking

M. W. G. Theunissen

Samen nascholen, maar hoe?

Garant, Leuven/Apeldoorn 1995

239 pagina's, f 49.90, ISBN 90-5350-409-5

De overheid streeft naar de opzet van Educatieve Faculteiten. Dat moeten organisaties worden waar lerarenopleidingen voor basis- en voortgezet onderwijs en schoolbegeleidingsdiensten nauw samenwerken. PABO's en Schoolbegeleidingsdiensten (SBD's) moeten samenwerken om leraren in basisscholen te ondersteunen bij hun taakuitvoering. In de praktijk blijkt deze samenwerking veelal de lossere vorm aan te nemen van een Educatieve Federatie. De bestaande vormen van samenwerking hebben vooral betrekking op nascholingsactiviteiten. Op dit terrein is van wederzijdse afhankelijkheid sprake tussen PABO's en SBD's door de wijze waarop de nascholing gefinancierd wordt en door de verdeling van deskundigheden over beide typen instellingen. Ten tijde van het promotie-onderzoek van Theunissen heeft de nascholing waarbij wordt samengewerkt met name betrekking op Onderwijskundig Management. De nascholingsactiviteiten binnen dit thema staan dan ook centraal in haar onderzoek.

De doelstelling van het onderzoek is het opsporen, beschrijven en evalueren van structurele samenwerkingsvormen tussen PABO's en SBD's bij nascholing. De probleemstelling is welke samenwerkingsvormen te onderscheiden zijn bij de samenwerking van PABO's en SBD's bij nascholing, en welke effecten de gekozen samenwerkingsvormen hebben op nascholing. De probleemstelling is uitgesplitst in een zestal onderzoeksvragen die betrekking hebben op structurele kenmerken van de samenwerking, problemen bij de ontwikkeling en organisatie van de samenwerking inclusief gekozen oplossingen, waardering van de samenwerking door diverse betrokkenen bij de nascholing, effecten van samenwerking op de nascholing, en tot slot ervaringen van cursisten met de samenwerking bij nascholing.

Voor haar theoretisch kader leunt Theunissen sterk op de netwerktheorie van Godfroy. Het betreft een theorie over structuren, strategieën en processen van samenwerking tussen interdependente organisaties in zich ontwikkelende inter-

organisationale netwerken. Veel interorganisatietoel onderzoek richt de aandacht vooral op de afzonderlijke organisaties, terwijl het in de theorie van Godfroy nu juist gaat om de samenwerking tussen organisaties. De kenmerken van het netwerk als geheel zijn niet zonder meer te reduceren tot kenmerken van afzonderlijke deelnemende instellingen. Dit maakt de theorie van Godfroy tot een aantrekkelijk uitgangspunt voor de studie van Theunissen.

De opzet van het onderzoek bestaat uit een survey en een case study. In het survey zijn alle nascholingscoördinatoren van SBD's en PABO's bevraagd door middel van een vragenlijst (respons 68%). De case-study is uitgevoerd in vier samenwerkingsverbanden die zijn geselecteerd op basis van de uitkomsten van het survey aangaande het samenwerkingsstype. Van deze samenwerkingsverbanden zijn het algemeen management van de instellingen, coördinatoren en uitvoerders bevraagd door middel van interviews (totaal 22). In de case-study zijn tevens de cursisten/afnemers van de nascholing bevraagd door middel van een vragenlijst (respons 46%).

Hoewel Godfroy in zijn netwerktheorie meerdere netwerktypen onderscheidt, vindt Theunissen in grote lijnen twee clusters van netwerkstructuren, waarbinnen verschillende typen een plaats hebben. In samenwerkingsverbanden met een praktische en inhoudelijke inslag wordt veel aandacht besteed aan uitwisseling van materialen en personeel, onderlinge afstemming en de ontwikkeling van een gezamenlijk produkt. In samenwerkingsverbanden met een bestuurlijke inslag ligt het accent meer op krachtenbundeling voor het verkrijgen van subsidie, en op het hanteren van verschillen in belangen en visies op beleidsniveau. De samenwerkingsverbanden zijn doorgaans federatief gestructureerd. Met betrekking tot bepaalde zaken kan het coördinerend orgaan zelfstandig opereren. Uitvoerende groepen beschikken over een relatief grote zelfstandigheid. Binnen de netwerken vertonen de instellingen doorgaans verschil in cultuur (PABO's zijn meer procesgericht en SBD's resultaatgericht). De samenwerking op de werkvloer voltrekt zich in de meeste gevallen naar tevredenheid, onafhankelijk van het netwerktype. Volgens deelnemers aan de interviews uit de meerwaarde van de samenwerking zich in verbeterde nascholingsprodukten. Deze meerwaarde heeft echter

een grens; ondersteuning van scholen en cursisten met het oog op toepassing van het geleerde in de eigen praktijk neemt door de samenwerking niet toe. De cursisten zijn tevreden over de nascholing. Het gegeven dat de cursus in samenwerking heeft plaatsgevonden is echter goeddeels aan hen voorbij gegaan.

In het onderzoek van Theunissen staat een belangrijke thematiek centraal. Het belang van samenwerking tussen onderwijsinstellingen zal de komende jaren immers alleen maar toenemen. Onderzoek naar organisaties in het onderwijs richt zich doorgaans op afzonderlijke scholen en instellingen. In theoretisch en praktisch opzicht biedt het onderzoek daarom nieuwe invalshoeken. Een aantrekkelijk punt is ook dat in de theoretische benadering van Godfroy een actor-gerichte benadering van organiseren centraal staat. Deze benadering biedt een welkome afwisseling op het overheersende structureel-functionalistische denken en het systeemdenken in het Nederlandse en internationale onderzoek naar schoolorganisaties. Hoeveel vorderingen maakt het onderzoek van Theunissen bij deze verschillende perspectieven?

De doelstelling van het onderzoek doet vermoeden dat bij de onderzoekster de belangstelling in praktische opbrengsten groter is geweest dan de belangstelling voor theorievorming op het terrein van netwerken van onderwijsorganisaties. Dit vermoeden wordt door de uitwerking van het onderzoek bevestigd. Het theoretisch kader voor het onderzoek wordt gebaseerd op de netwerktheorie van Godfroy. In de uitwerking wordt echter naar verhouding sterk de nadruk gelegd op formele structuur- en functiekennmerken van de netwerken en de daarin deelnemende SBD's en PABO's. De uitwerking van de actor-gerichte dynamische aspecten van de theorie krijgt in het onderzoek slechts beperkt aandacht. De clustering van een aantal netwerkstructuurtypen tot twee hoofdclusters is een wat magere theoretische opbrengst. De conclusie dat netwerken in de tijd een ontwikkeling doormaken van beleidsgericht naar inhoudelijk gericht lijkt voorbarig vanwege het cross-sectioneel karakter van het onderzoek. Het gegeven dat oudere netwerken inhoudelijk gericht en jongere netwerken beleidsgericht zijn kan evengoed of wellicht beter verklaard worden door het tijdstip van oprichting van de netwerken in samenhang met

veranderingen in de door de auteur beschreven landelijke beleidscontext. In deze context wordt samenwerken met de dag een nijpender noodzaak om te overleven. Deze context plaatst opportunistische motieven om samen te werken (jonge, beleidsgerichte netwerken) op de voorgrond. Inhoudelijke motieven om samen te werken (oude, inhoudelijk gerichte netwerken) verdwijnen meer en meer naar de achtergrond. Naast conceptuele beperkingen staan ook methodische problemen een grotere theoretische opbrengst in de weg. Zo kan de waarde van een aantal correlaties tussen de variabele netwerktype en andere variabelen niet geschat worden, omdat deze variabelen eerder een rol spelen in de constructie van de netwerktypen. Het siert de auteur dat zij deze methodisch minder sterke plek zelf als probleem heeft opgevoerd.

De opbrengst van het onderzoek moet vooral op het evaluatieve en praktische vlak worden gezocht. Het onderzoek biedt een aardig inzicht in de stand van zaken van de vorming van Educatieve Federaties en vergelijkbare samenwerkingsverbanden in Nederland. Een interessante bevinding is dat de produkten waarbij SBD's en PABO's samenwerken relatief veel aftrek vinden, terwijl tegelijkertijd deze instellingen relatief weinig investeren in deze produkten en de afnemers van de produkten in de scholen de samenwerking niet als zodanig waarnemen. Ook is de constatering belangrijk dat een potentiële meeropbrengst van samenwerking - verbetering van de begeleiding bij de transfer van de training in de scholen - niet wordt gerealiseerd. Hier liggen aangrijpingspunten om op strategisch niveau de samenwerking nader te profileren, en om de samenwerking inhoudelijk verder uit te bouwen. In het evaluatieve deel van het onderzoek is een methodische onduidelijkheid waarom de structuur van het schriftelijke evaluatie-instrument voor scholen met een factoranalyse wordt bepaald. Het gaat hier niet om het opsporen van theoretische constructen, maar om heel praktische factoren zoals de bruikbaarheid van de cursus Onderwijskundig Management op school, en de tevredenheid van de cursisten over andere nascholingscursussen. De uiteindelijke categorieën van items zouden op basis van de tevens uitgevoerde betrouwbaarheidsanalyse toch wel naar voren zijn gekomen.

Een zinvolle maar in diepgang beperkte bevinding is dat de netwerktheorie in aanzet bruik-

baar is om samenwerking tussen onderwijsinstellingen in kaart te brengen. Resumerend kan worden vastgesteld dat de in leesbare stijl geschreven dissertatie vooral interessant is voor lezers met belangstelling voor empirisch onderzoek naar de praktische kanten van samenwerking tussen onderwijsinstellingen, met name samenwerking tussen PABO's en SBD's.

J. Imants

Mededelingen

Promoties

Universiteit Utrecht

Promovendus: Drs. J. H. M. Castelijns

Titel proefschrift: Beelden van bekwaamheid.

Een onderzoek naar de implementatie en effecten van het programma 'Responsive instructie' voor leerkrachten van groep 2 van de basisschool

Promotor: Prof. Dr. L. M. Stevens

Co-promotor: Dr. K. M. Stokking

Datum: 10 juni 1996

Universiteit Utrecht

Promovendus: Drs. M. C. Luna Rubio

Titel proefschrift: Leerlingbegeleiding: wens en werkelijkheid. Een studie naar de activiteiten van de leerlingbegeleiding

Promotores: Prof. Dr. N. Deen en

Prof. Dr. H. 't Hart

Datum: 19 juni 1996

Inhoud andere tijdschriften

Comenius

16e jaargang, nr. 2, 1996

Het effect van ouderlijke ruzie en echtscheiding op het welzijn van middelbare scholieren, door J. Dronkers

Verheffing door handenarbeid, door J. Jongkind
Kansen, plichten en het marktgebod. Intercultureel management en inburgeringsprogramma's als interventies in interetnische verhoudingen, door F. Glastra en P. Schedler

Een sterk plekbeseft: Dennendal revisited, door E. Tonkens en I. Weijers

Angst voor verlies van vrouwelijkheid. Honderd jaar discussie over de relatie tussen sekse en onderwijs en wetenschap, door S. Grotenhuis

Pedagogisch Tijdschrift

21e jaargang, nr. 2, 1996

Over het wezen en de betekenis van de koloniaal-pedagogische traditie in België en Nederland, door A. T. Boone en M. Depaepe
Onderwijs en opvoeding in De Nederlandse koloniën, door A. T. Boone
Koloniale pedagogiek in Nederlands-Indië, door

J. E. A. M. Lelyveld

De bijzondere situatie van Europese kinderen in de voormalige Oost-Indische kolonie, door G. C. Brantas

Onderwijs en missie strategie (1872-1940). Een case-study over de Chinamissies, door C. Dujardin

'Rien ne va plus ...'. De ondergang van de koloniale educatieve structuren in Zaïre (1960-1995), door M. Depaepe

Tijdschrift voor Orthopedagogiek

35e jaargang, nr. 5, 1996

Rekenen en rekenproblemen, door J. E. H. van Luit en A. J. J. M. Ruijsse-naars

Rekenvaardigheden van kleuters (1). Onderzoek in het basisonderwijs, door B. A. M. van de Rijt, J. E. H. van Luit en A. H. Pennings

Rekenvaardigheden van kleuters (2). Onderzoek bij In hun Ontwikkeling Bedreigde Kleuters, door E. A. M. Schopman en J. E. H. van Luit
Een orthodidactische aanpak van rekenproblemen bij rekenzwakke leerlingen in het regulier basisonderwijs, door P. Ghesquière, A. Ruijsse-naars, H. Grietens en E. Luyckx

Tijdschrift voor Orthopedagogiek

35e jaargang, nr. 6, 1996

Traumabeleving door geadopteerden, door R. A. C. Hoksbergen

Rekenproblemen: de andere kant van de medaille, door G. M. van der Aalsvoort

Twee provinciale plannen jeugdhulpverlening gezien in samenhang met het Rijksplan, door J. H. L. Oud en J. J. M. van der Eerden

Ontvangen boeken

Braster, J. F. A., *De identiteit van het openbaar onderwijs* (dissertatie). Wolters-Noordhoff, Groningen, 1996, f 42,50.

Kats, E., *Metaforen en methodologieën. Ontwikkelingen in praktijkgericht onderzoek*. Wolters-Noordhoff, Groningen, 1996, f 49,50.

Lockhorst, D., & Meer, T. van der, *Tussen luisterschool en werkplaats. Handleiding voor creatieve leerplanontwikkeling en begeleiding*. Wolters-Noordhoff, Groningen, 1996, f 49,50.

Rispens, J., Hermanns, J. M. A., & Meeus, W. H. J. (red.), *Opvoeden in Nederland*. Van Gorcum, Assen, 1996, f 42,50.

3-75 OCT 6 1996

Inhoudsopgave

Pedagogische Studiën Jaargang 73, nummer 5/1996

L. Verschaffel en E. De Corte

Leren realistisch modelleren en interpreteren van vraagstukken.
Een exploratief onderwijsexperiment bij leerlingen van de boven-
bouw van de basisschool 322

F. W. P. van der Slik, J. R. M. Gerris en A. J. A. Felling

Opleiding van ouders en de opvoeding van hun kinderen. Een on-
derzoek in de Kohn-traditie 338

S. Veenman en J. Raemaekers

Retentie-effecten van een nascholingsprogramma voor effectieve
instructie en klasmanagement 357

Kroniek

N. Verloop en L. Verschaffel (red.)

Het AERA-Congres 1996, New York 372

Boekbesprekingen 395

BIBLIOTHEEK DER
RIJKSUNIVERSITEIT
UTRECHT

Leren realistisch modelleren en interpreteren van vraagstukken. Een exploratief onderwijsexperiment bij leerlingen van de bovenbouw van de basisschool

L. Verschaffel en E. De Corte*

Samenvatting

Recent onderzoek heeft aangetoond dat leerlingen uit de bovenbouw van de basisschool een sterke neiging hebben om hun ervaringskennis omtrent de concrete, reële situatie of context waarop een rekenvraagstuk betrekking heeft te negeren tijdens het oplossingsproces. In onderhavig onderwijsexperiment werd nagegaan of het mogelijk is om betekenisvolle positieve leer-effecten te bereiken op het vlak van het realistisch modelleren en interpreteren van wiskundige toepassingsproblemen. Daartoe werden leerlingen ondergedompeld in een onderwijsleeromgeving waarin vraagstukken juist opgevat werden als oefeningen in realistisch modelleren en interpreteren, eerder dan als oefenstof voor het indrillen van stereotiepe, formele rekenoperaties. De leergang werd uitgevoerd in een vijfde leerjaar van een Vlaamse basisschool en bestond uit vijf leereenheden van ongeveer 2,5 uur verspreid over 2,5 weken. Intussen volgden de leerlingen van twee controlegroepen gewoon reken/wiskundeonderwijs. Naast specifieke leerresultaten, peilden de toetsen naar transfer en retentie van het geleerde. De onder-

zoeksresultaten leveren steun voor de hypothese dat het inderdaad mogelijk is om de dispositie tot realistisch modelleren en interpreteren van leerlingen uit de bovenbouw van de basisschool op betekenisvolle en positieve wijze te beïnvloeden.

1 Inleiding

Vraagstukken vormen een belangrijk onderdeel van het reken/wiskundeprogramma van de basisschool. Eén van de voornaamste redenen om dit soort van problemen in de reken/wiskundelessen in te schakelen, is ervoor te zorgen dat de leerlingen de wiskundige begrippen en rekentechnieken die ze tijdens deze lessen verworven hebben, kunnen toepassen in levens-echte of levensnabije (probleem)situaties (De Corte & Verschaffel, 1989; Treffers & De Moor, 1990; Verschaffel & Gravemeijer, 1990). Recent onderzoek heeft evenwel aangetoond dat het vraagstukkenonderwijs bij leerlingen de neiging doet ontstaan om hun alledaagse ervaringskennis en hun realistische overwegingen omtrent de concrete situatie of context waarop het vraagstuk betrekking heeft, te negeren tijdens de verschillende fasen van het oplossingsproces, nl. de opbouw van de initiële probleemrepresentatie, de constructie van een passend wiskundig model, de uitvoering van de rekenkundige operatie(s) die in dit model vervat zit(ten), en de interpretatie en evaluatie van de uitkomst van het rekenwerk (Davis, 1989; De Corte & Verschaffel, 1985; Goffree, 1983; Greer, 1993; Kilpatrick, 1987; Nelissen, 1977; Nesher, 1980; Reusser, 1988; Schoenfeld, 1991; Silver, Shapiro & Deutsch, 1993; Treffers & De Moor, 1990; Verschaffel & De Corte, ter perse). Dit zou één van de voornaamste verklaringen zijn waarom de traditionele schoolvraagstukkendidactiek er onvol-

* De auteurs bedanken de directeur en de leerkrachten van de basisschool waarin het onderzoek plaatsvond voor de bereidwillige en vlotte medewerking. Een bijzonder woord van dank gaat uit naar G. Bonne, leerkracht van de experimentele klas, voor zijn vele opmerkingen en suggesties bij voorlopige versies van het trainings- en toetsmateriaal en voor zijn hulp bij de realisatie en registratie van de experimentele leergang. Tevens bedanken we A. Treffers, K. Gravemeijer en D. van Eerde voor hun commentaren op een eerste concept van het onderzoek. Tot slot ook een woord van dank aan R. Stroobants voor zijn hulp bij de statistische verwerking van de gegevens.

doende in slaagt om bij leerlingen deze bekwaamheid in het oplossen van reële toepassingsproblemen te ontwikkelen.

In twee recente studies verzamelden Verschaffel, De Corte en Lasure (1994; ter perse) overtuigende empirische steun voor deze bewering. In een eerste onderzoek kregen 75 vijfdeklassers afkomstig uit drie verschillende Vlaamse scholen tijdens een gewone rekenles een schriftelijke toets voorgelegd bestaande uit 10 itemparen. Het ene vraagstuk uit elk opgavenpaar was een standaardopgave (S-item) die 'probleemloos' gemodelleerd en opgelost kan worden door middel van één of twee rekenkundige basisbewerkingen met de getallen uit de opgave (bijv. "Wim heeft 4 planken van 2 meter gekocht. Hoeveel planken van 1 meter kan hij daaruit zagen?"). Bij de andere opgave uit elk paar vormt de omzetting in een passend wiskundig model wél een probleem (P-item), tenminste als men de context waarrond deze opgave is opgebouwd 'au sérieux' neemt (bijv. "Wim heeft 4 planken van 2,5 meter gekocht. Hoeveel planken van 1 meter kan Wim daarvan zagen?"). Het aantal antwoorden en/of commentaren op de P-items waaruit bleek dat de oplosser rekening hield met de reële situatie uit de betreffende vraagstukken, was alarmerend laag: slechts 17% van alle leerlingreacties op deze P-items was als 'realistisch' te bestempelen. Bij het zoëven vermelde planken-vraagstuk bijvoorbeeld gaf slechts 10% van de leerlingen te kennen dat men in werkelijkheid slechts 2 planken van 1 meter kan zagen uit één plank van 2,5 meter, en dat derhalve 4 van dergelijke planken slechts 8 planken van 1 meter kunnen opleveren in plaats van 10. De overige leerlingen reageerden vrijwel allemaal met het verwachte, stereotiepe antwoord " $4 \times 2,5 = 10$ " zonder enige commentaar. In een tweede studie constateerden Verschaffel e.a. (ter perse) dat deze neiging van leerlingen tot niet-realistisch reageren erg weerbarstig is, in die zin dat leerlingen die een P-item aanvankelijk op niet-realistische wijze hadden beantwoord, over het algemeen niet geneigd waren dit niet-realistisch antwoord in te ruilen voor een (meer) realistisch, zelfs na twee sterke 'hints' vanwege de interviewer in deze richting.

De studies van Verschaffel e.a. (1994; ter perse) tonen dus aan dat leerlingen na verschillende jaren klassiek vraagstukkenonderwijs

geleerd hebben om bij het omzetten van een wiskundig toepassingsprobleem in een wiskundige berekening of bij de interpretatie van de uitkomst van dit rekenwerk, weinig of geen ruimte te laten voor realistische ervaringskennis en praktische overwegingen. Deze studies verschaffen ons evenwel geen inzicht in de wijze waarop dit gebrek aan 'realiteitszin' tijdens het oplossen van wiskundige toepassingsproblemen bij leerlingen tot ontwikkeling komt. Verschaffel e.a. (1994; ter perse) spreken het vermoeden uit dat dit het resultaat is van de wisselwerking tussen

- 1 het stereotiep en levensvreemd karakter van de overgrote meerderheid van de rekenvraagstukken waarmee leerlingen (onder meer via de rekenboeken) in de vraagstukkenlessen worden geconfronteerd, en
- 2 de manier waarop er tijdens deze lessen door de leerkracht met deze vraagstukken wordt omgesprongen.

Door toedoen van deze aspecten van de klas-sieke vraagstukkendidactiek zou bij leerlingen langzaam maar zeker de dispositie groeien om de concrete situatie waarin een wiskundig toepassingsprobleem ingebed zit, niet al te ernstig te nemen en om de eigen concrete ervaringskennis daaromtrent te negeren. Er is echter weinig empirische evidentie voor deze hypothetische verklaring.

In dit artikel wordt verslag uitgebracht van een exploratief onderzoek waarin het effect is nagegaan van een experimenteel (remediërings)programma¹ op de dispositie tot realistisch modelleren en interpreteren van leerlingen uit de bovenbouw van de basisschool. Centraal in dit programma staat het ontwikkelen van de geneigdheid en de vaardigheid tot het inschakelen van ervaringskennis en van realistische overwegingen in de diverse fasen van het oplossingsproces, in het bijzonder

- 1 tijdens de constructie van een gepast wiskundig model voor een gegeven probleem-situatie (= realistisch modelleren), en
- 2 bij het interpreteren van de uitkomst van het rekenwerk in functie van de oorspronkelijke vraagstelling (= realistisch interpreteren).

2 Opzet

2.1 Overzicht van de onderzoeksopzet

Het onderzoek vond plaats gedurende het eerste trimester van het schooljaar 1994-95. Een experimentele klas (E-klas) en twee controleklassen (C1 en C2) waren erin betrokken.

De lessen in de E-klas werden grotendeels verzorgd door de eerste auteur. Dit gebeurde echter in nauw overleg met de 'gewone' leerkracht, die zelf ook enkele kleinere onderdelen voor z'n rekening nam. In de twee controleklassen werd het gewone reken/wiskunde-programma afgewerkt. De experimentele leergang bestond uit vijf thematische leereenheden (LE) van 2 à 2,5 uur verspreid over 2,5 weken. Bij het begin en aan het einde van de leergang werden in de E-klas en in de twee controleklassen schriftelijke voor- en natoetsen afgenomen waarmee de vaardigheid van de leerlingen in het realistisch modelleren en interpreteren van context-gebonden wiskundige toepassingsproblemen werd nagegaan. De leerlingen uit de eerste controlegroep kregen net voor de natoets een korte introductie over realistisch modelleren en interpreteren (zie verder). Een maand na het beëindigen van de leergang werd er in de E-klas nog een retentietoets afgenomen. Informatie over het concrete verloop van het experimenteel programma werd verkregen via de analyse van video-opnamen van alle lessen én van al het schriftelijk leerlingmateriaal.

2.2 Subjecten

De leerlingen uit de E-klas waren vijfdeklassers, terwijl die uit de twee controleklassen zesdeklassers waren. Deze drie klassen maakten deel uit van eenzelfde lagere jongensschool en telden respectievelijk 19, 18 en 17 leerlingen. Deze school kan getypeerd worden als een dorpsschool die haar leerlingen recruteert uit alle lagen van de plaatselijke bevolking.

In deze school wordt sedert meerdere jaren gebruik gemaakt van een bekende traditionele Vlaamse rekenmethode, nl. "Reken Raak". Uit voorafgaande gesprekken met de directie en de leerkrachten was gebleken dat er in deze school tot dan toe geen extra aandacht besteed was aan het leren realistisch modelleren en interpreteren van wiskundige toepassingsituaties. Deze gesprekken brachten ook nog aan het licht dat men tot dan toe bij het reken/wiskunde-onder-

wijs weinig of geen gebruik gemaakt had van groepswork.

2.3 Beschrijving van het experimenteel programma

2.3.1 Opgavenaanbod

Een eerste belangrijke pijler van het programma betreft de aard van het opgavenaanbod. De traditionele standaardopgaven werden vervangen door een reeks problemen die stuk voor stuk speciaal ontworpen waren om leerlingen te confronteren met het probleem van realistisch modelleren en interpreteren van contextgebonden toepassingsituaties. Elk van de vijf leereenheden (LE) was opgebouwd rond één welbepaald aspect van realistisch modelleren en interpreteren.

De eerste LE ging over het leren interpreteren van het resultaat van een niet-opgaande deling. Deze LE begon met een verhaal over een regiment van 300 soldaten die diverse militaire activiteiten uitvoerden. Bij elke militaire activiteit hoorde een vraag, die altijd via dezelfde rekenkundige bewerking kon worden opgelost, maar telkens een ander antwoord vereiste (zie Bijlage).

LE 2 handelde over additieve probleemsituaties rond sets die gemeenschappelijke elementen hebben en derhalve niet op te lossen zijn door de aantallen van de diverse sets gewoon bij elkaar op te tellen of van elkaar af te trekken. Het uitgangsprobleem betrof een jongen die strips van "Suske en Wiske" verzamelde en ter gelegenheid van zijn verjaardag een groot pak oude albums van "Suske en Wiske" kreeg waaronder verscheidene albums die hij reeds bezat. De leerlingen dienden te berekenen hoeveel strips er na dit geschenk nog ontbraken in de verzameling van deze jongen.

In de derde LE stonden probleemsituaties centraal waarbij het niet duidelijk was of de uitkomst van een rekenkundige operatie met de twee gegeven getallen het exacte antwoord opleverde dan wel een getal dat 1 groter of 1 kleiner is dan het juiste antwoord. In het uitgangsprobleem dat bij deze LE hoorde kregen de leerlingen de nummers van het eerst en het laatst verkochte toegangsbiljet op een bepaalde dag in het plaatselijk zwembad, en werd er gevraagd hoeveel kaartjes er op deze dag verkocht waren.

Het uitgangsverhaal van de vierde LE ging over een jongen die een schommel wilde maken aan een 5 meter hoge tak van een boom; gevraagd werd hoeveel meter touw deze jongen daarvoor nodig had. In deze LE leerden de leerlingen dat er bij het opstellen van een wiskundig model en/of bij de interpretatie van de uitkomst van het rekenwerk geregeld gebruik moet worden gemaakt van informatie-elementen die weliswaar niet expliciet of onmiddellijk 'gegeven' zijn maar wel (eenvoudig) te achterhalen zijn door beroep te doen op eigen ervaringskennis.

In LE 5 tenslotte leerden de leerlingen onderscheid maken tussen toepassingsituaties die wel en niet adequaat gemodelleerd kunnen worden in termen van lineaire proportie (a:b zoals c:d). Als uitgangsprobleem fungeerde een verhaal over een jonge atleet waarvan de recordtijd op de 100 meter hardlopen gegeven was en de recordtijd op de 400 meter moest worden voorspeld.

Naast het uitgangsprobleem kwamen in elke LE ook nog verscheidene andere contextgebonden toepassingsopgaven voor waaraan dezelfde of een analoge modellerings- of interpretatiemoeilijkheid ten grondslag lag (zie paragraaf 2.3.2).

2.3.2 Instructiemethoden

Niet alleen het opgavenaanbod maar ook de instructietechnieken die in het programma werden gebruikt verschilden aanzienlijk van die uit een klassieke vraagstukkenles (voor een typering van het klassieke vraagstukkenonderwijs, zie o.m. De Corte & Verschaffel, 1989).

Zoals gezegd in paragraaf 2.3.1, werd telkens uitgegaan van een aantrekkelijk, uitdagend uitgangsprobleem waarbij er zich een bepaalde modellerings- of interpretatiemoeilijkheid voordeed. Daaraan werd gedurende 15 à 20 minuten gewerkt in zes vaste heterogene groepen van drie à vier leerlingen. De leerlingen van elk groepje werden aangespoord om het probleem eerst individueel op een klein (klad)blad (A4-formaat) op te lossen en pas daarna uit de individuele oplossingen één gezamenlijk antwoord te distilleren en dit op een groot blad (A3-formaat) te noteren. Op deze bladen moest overigens niet enkel het antwoord genoteerd worden, maar tevens de gevolgde oplossingsweg, evenals een reflectie op

het oplossingsproces en op eventuele moeilijkheden die zich gedurende dit proces hadden voorgedaan (zie Bijlage 1). Tijdens het groepswork liep de lesgever rond om zo goed mogelijk zicht te krijgen op de wijze waarop het probleem in de diverse groepjes werd aangepakt en opgelost. Tevens beantwoordde hij eventuele vragen en gaf hij soms strategische aanwijzingen en hulp (bijv. "Lees de opgave nog eens heel aandachtig", "Maak eens een schets van de probleemsituatie", "Hebben jullie dit antwoord gecontroleerd?").

Na dit groepswork werden de verschillende antwoorden op en oplossingswijzen van het uitgangsprobleem evenals de moeilijkheden die men daarbij had ondervonden, gedurende 15 à 20 minuten klassikaal besproken. Aan het einde van deze klassikale bespreking nodigde de lesgever de leerlingen uit om wat ze op grond van de oplossing en de bespreking van het uitgangsprobleem hadden geleerd samen te vatten, en trachtte hij het geleerde kernachtig in meer algemene termen te (her)formuleren.

Vervolgens kregen de leerlingen een werkblad waarop vier nieuwe toepassingsproblemen stonden, die wederom gedurende 15 à 20 minuten in groepjes werden opgelost. Aan sommige van deze opgaven lag hetzelfde modellerings- of interpretatieprobleem ten grondslag als aan het uitgangsprobleem; daarnaast bevatte dit werkblad echter ook één of twee 'probleemloze' standaardopgaven (S-items).

Daarop volgde opnieuw een klassikale bespreking (15 à 20 minuten), die op dezelfde manier verliep als de klassikale discussie over het uitgangsprobleem.

Bij elke LE hoorde er tenslotte ook nog een individuele opdracht, waarin de leerlingen nogmaals geconfronteerd werden met een toepassings situatie met dezelfde modellerings- of interpretatiemoeilijkheid als het uitgangsprobleem. Ook deze individuele opdracht werd achteraf klassikaal besproken.

Tijdens het groepswork werd zoals gezegd gewerkt in zes vaste, heterogene groepen van drie of vier leerlingen. Bij de verdeling van de 19 leerlingen over deze zes groepen werd ervoor gezorgd dat elke groep één 'sterke', één of twee 'middelmatische' en één 'zwakke' leerling telde. Voor de indeling van de leerlingen volgens rekenniveau werd een beroep gedaan op het oordeel van de leerkracht.

Bij wijze van illustratie zijn in bijlage de werkbladen opgenomen waarvan gebruik gemaakt is in LE 1. Werkblad 1 en 2 bevatten het uitgangsprobleem met de bijhorende vragen, werkblad 3 de reeks met oefenopgaven en werkblad 4 de individuele opdracht waarmee LE 1 afgesloten werd. De werkbladen van de overige leereenheden zijn te vinden in Verschaffel (1995).

2.3.3 *Klassecultuur*

Een derde belangrijke pijler van het programma was de expliciete gerichtheid op de ontwikkeling van een waardevolle wiskundige dispositie bij de leerlingen, via het creëren van een nieuwe klassecultuur waarin vraagstukkenlessen niet opgevat worden als trainingen in het routine-matig gebruik van voorgeschre-

ven quasi-algoritmische oplossingsmethoden, maar als een proefterrein voor het leren zelfstandig en kritisch gebruiken van de verworven wiskundige kennis en vaardigheden in levens-echte toepassingsituaties. Dit werd o.m. bewerkstelligd via het expliciet bespreken van en negotiëren over spelregels en normen omtrent:

- wat er verstaan wordt onder een goed wiskundig toepassingsprobleem (bijv. "sommige problemen kunnen op meerdere manieren (correct) beantwoord worden"), een goede oplossing (bijv. "soms is het zinvoller om geen exact antwoord te geven maar een schatting te maken") of een goede oplossingsmethode (bijv. "ook ervaren rekenaars tellen soms nog wel eens op hun vingers");
- de rol van de leerkracht en van de leerlingen in een reken/wiskundeles (bijv. "jullie moe-

Tabel 1

Tien P-items uit de voortoets

1A*	1180 supporters moeten per bus naar een voetbalwedstrijd vervoerd worden. Hoeveel bussen zijn er nodig, als je weet dat er maximum 48 supporters in één bus kunnen?
1B	Een groep van 228 toeristen wil vanop het dak van het hoogste flatgebouw van de stad genieten van een panoramisch uitzicht. In het flatgebouw is er slechts één lift die tot het dak gaat. Het maximum aantal personen is 24. Hoeveel keer zal de lift omhoog moeten alvorens al deze toeristen boven zijn?
2A	Op het einde van het schooljaar trachten de 66 leerlingen van het derde leerjaar hun zwemdiploma te behalen. Om dit diploma te behalen, moet je 100 meter ver kunnen zwemmen en 1 minuut lang kunnen watertrappen. 13 leerlingen slagen er niet in om 100 meter ver te zwemmen en 11 leerlingen slagen er niet in om één minuut lang te watertrappen. Hoeveel leerlingen van het derde leerjaar krijgen er een zwemdiploma?
2B	Karel en Joris zitten in dezelfde klas en verjaren op precies dezelfde dag. Karel heeft 9 vrienden die hij wil uitnodigen op z'n verjaardagsfeest en Joris heeft er 12. Karel en Joris besluiten om samen één verjaardagsfeest te geven en ze nodigen al hun vrienden uit. Alle vrienden komen. Hoeveel vrienden zijn er op het feest?
3A	Enige tijd geleden werd op school het afscheid van de directeur gevierd. Hij was zonder onderbreking directeur van 1 januari 1959 tot 31 december 1993. Hoeveel jaar is hij directeur geweest?
3B	Dit jaar werd het rock-concert "Torhout-Werchter" voor de vijftiende keer gehouden. In welk jaar werd dit concert voor het eerst georganiseerd?
4A	De afstand tussen twee palen bedraagt 12 meter. Hoeveel stukken touw van 1,5 meter lang moet men aan elkaar knopen om de beide palen met elkaar te verbinden?
4B	Steven heeft 8 planken van 2,5 meter gekocht. Hoeveel planken van 1 meter kan Steven hiervan zagen?
5A	De recordtijd van Sven op 50 meter schoolslag is 54 seconden. In hoeveel tijd zwemt hij de 200 meter schoolslag?
5B	Deze vaas wordt gevuld met water. Na 10 seconden staat het water 3,5 cm hoog. Hoe hoog zal het water staan na 30 seconden? (Deze opgave was vergezeld van een tekening van een kegelvormige vaas)

* De getallen verwijzen naar de LE uit de experimentele leergang waarin dat soort van modellerings- en interpretatiemoelijkheid centraal stond. Het A-item van elk paar is het leeritem; het B-item het nabije-transferitem.

ten niet verwachten dat ik alleen zal bepalen welke oplossing de beste is; dat is iets waarover we samen zullen overleggen") (zie ook Cobb, 1989; Gravemeijer, 1994).

2.4 Toetsen

2.4.1 Voortoets

In de drie klassen (E, C1 en C2) werd enkele dagen voor het begin van de experimentele leergang een collectieve schriftelijke voortoets afgenomen. Deze toets bestond uit 15 vraagstukken: vijf paren van twee 'problematische' opgaven (P-items) (één paar per LE) en vijf gewone standaardopgaven (S-items) die als bufferitems fungeerden. Tabel 1 geeft een overzicht van de 10 P-items uit de voortoets.

Het verschil tussen de beide P-items van elk paar lag hierin dat het eerste opgebouwd was rond een situatie of context die sterk verwant was aan één van de probleemsituaties waarmee de leerlingen tijdens de overeenkomstige LE hadden gewerkt (= het leeritem), terwijl het andere opgebouwd was rond een situatie of context die minder gelijkenis vertoonde met één van de P-items uit deze LE (= nabijef-transferitem).

Van deze toets werden twee verschillende versies gemaakt, die onderling verschilden wat betreft de volgorde waarin de opgaven gepresenteerd werden.

De instructies bleven beperkt tot een minimum. De opdracht luidde eenvoudigweg om voor elk vraagstuk het antwoord neer te schrijven in het daartoe voorziene 'antwoord-vak', en om in het 'uitleg-vak' daaronder te noteren hoe men aan dit antwoord was gekomen of om uit te leggen waarom men het betreffende vraagstuk niet (goed) kon oplossen.

2.4.2 Natoets

Enkele dagen na het beëindigen van de experimentele leergang werd in de experimentele klas een collectieve natoets afgenomen, die als een parallelversie van de voortoets te beschouwen is. Dit betekent dat ook deze toets bestond uit vijf paren van P-items (één paar per LE) en vijf S-opgaven die als bufferitems fungeerden.

Zoals bij de voortoets, werden er van de natoets twee versies opgesteld die onderling verschilden wat betreft de volgorde van de opgaven.

De leerlingen uit de twee controleklassen kregen dezelfde natoets als de leerlingen van de E-klas. In controleklas 2 werd de natoets - net zoals in de E-klas - zonder speciale introductie gegeven. In controleklas 1 daarentegen werd vooraf gedurende 10 à 15 minuten een introductie gegeven waarin de leerlingen uitgelegd werd dat routine-matige antwoorden op vraagstukken soms inadequaet zijn vanuit realistisch oogpunt. Tevens werden enkele voorbeelden van dergelijke inadequate antwoorden gegeven (bijv. het vraagstuk "Er zitten 10 vogels in een boom. Een jager komt en schiet er 2 neer. Hoeveel vogels zitten nog in deze boom?" beantwoorden met "8 vogels" in plaats van met "geen enkele meer"). Aan het einde van deze introductie werden de leerlingen er uitdrukkelijk voor gewaarschuwd dat de natoets verscheidene opgaven bevatte waarvoor dergelijke routine-matige antwoorden inadequaet zijn. Door de eventuele leerwinst van de C1-leerlingen tussen voor- en natoets te vergelijken met die van de E- en de C2-groep, zou kunnen worden nagegaan of een dergelijke beperkte ingreep volstaat om het aantal niet-realistische oplossingen bij context-gebonden wiskundige toepassingen te doen afnemen.

2.4.3 Retentietoets

Tenslotte werd de leerlingen uit de E-groep een maand na het beëindigen van het programma ook nog een retentietoets afgenomen.² Deze afname vond plaats in het kader van een gewone reken/wiskundeles én in afwezigheid van de onderzoekers. Verder werd er bij het klaarmaken van de retentietoets voor gezorgd dat deze toets qua uiterlijke kenmerken (soort van papier, bladschikking, lettertype...) zoveel mogelijk verschilden van de voor- en natoets.

De retentietoets bestond eveneens uit 15 opgaven: 10 P-items en vijf S-opgaven die als bufferitems fungeerden. Vijf van de 10 P-items waren parallel-items van de vijf nabijef-transferopgaven uit de voor- en de natoets (zie Tabel 1). De vijf resterende P-items waren eveneens 'problematische' vraagstukken waarvan de relatie met de modellerings- en interpretatiemoeilijkheden uit het experimenteel programma evenwel minder duidelijk was (= verre-transferitems), maar waarvan we uit ander onderzoek wisten dat leerlingen die traditioneel reken/wiskunde-onderwijs achter de

rug hebben er zelden of nooit in realistische zin op reageren (Greer, 1993; Verschaffel e.a., 1994). Twee voorbeelden van deze laatste groep van vijf verre-transferopgaven uit de retentietoets zijn "Bruno en Saskia gaan naar dezelfde school. Bruno woont 17 kilometer van school vandaan en Saskia 8 kilometer. Hoever wonen Bruno en Saskia van elkaar?" en "Rob is geboren in 1982. Het is nu 1994. Hoe oud is Rob?".

Zoals in de voor- en natoets, werden de opgaven uit de retentietoets in twee verschillende volgorden aangeboden.

2.5 Analyse van de gegevens

2.5.1 Scoring van de reacties op voor-, na- en retentietoets

Alle reacties van de leerlingen op de P-items van de voortoets, de natoets en de retentietoets werden op de volgende twee punten geanalyseerd: (1) het antwoord dat de leerling in het daartoe voorziene 'antwoord-vak' genoteerd had, en (2) eventuele rekenkundige bewerkingen en/of opmerkingen in het 'uitleg-vak'.

De analyse van wat de leerling in het 'antwoord-vak' genoteerd had was erop gericht onderscheid te maken tussen 'realistische' en 'niet-realistische' antwoorden. Een realistisch antwoord (RA) is een antwoord dat voortvloeit uit het effectief gebruik van ervaringskennis en realistische overwegingen in verband met de context waarrond het vraagstuk is opgebouwd. Bijvoorbeeld: het meest aangegeven realistisch antwoord op het planken-vraagstuk uit paragraaf 1 is '8 planken'. Immers, zoals eerder vermeld, kan men 'in werkelijkheid' uit elke plank van 2,5 meter slechts 2 planken van 1 meter zagen, hetgeen impliceert dat er uit 4 van deze planken slechts 8 planken van 1 meter kunnen worden gezaagd. Het niet-realistisch antwoord op dit P-item is '10 planken', - het resultaat van de voor-de-hand-liggende bewerking met de twee gegeven getallen, nl. $4 \times 2,5 =$.

Naast het antwoord van de leerling dat in het 'antwoord-vak' genoteerd stond, werden zoals gezegd ook de eventuele berekeningen en commentaren in het 'uitleg-vak' bekeken. Meer bepaald werd nagegaan of in deze notities enig spoor te vinden was van eventuele moeilijkheden die de leerling ondervond had bij het

wiskundig modelleren van het vraagstuk en/of bij het interpreteren van de uitkomst van het rekenwerk, als gevolg van de activering van realistische overwegingen. Een voorbeeld van zo'n opmerking is de commentaar "Maar dan moet Wim de overblijvende stukken van 0,5 meter wel stevig aan elkaar lijmen" in het 'uitleg-vak' van een leerling die het planken-vraagstuk met '10' beantwoord heeft. Wanneer er in het 'uitleg-vak' enig spoor te vinden was van activering van een dergelijke realistische overweging, werd de reactie van de leerling op het betreffend P-item eveneens beschouwd als 'realistisch'.

De uitdrukking 'realistische reactie' (RR) slaat derhalve zowel op een realistisch antwoord (RA) als op een antwoord dat niet als 'realistisch' gescoord werd maar dat wél gepaard ging met één of andere realistische commentaar in het uitleg-vak.³ De uitdrukking 'niet-realistische reactie' (NR) daarentegen verwijst naar een reactie waarbij noch in het antwoord noch in de uitleg van een leerling sprake was van activering van relevante, context-gebonden ervaringskennis.

2.5.2 Analyse van het verloop van de leergang

Zoals gezegd, wilden we met deze studie niet alleen het effect nagaan van de ontwikkelde leergang op de dispositie en de vaardigheid van de leerlingen in het realistisch modelleren van wiskundige toepassingsproblemen. Een bijkomend doel was dieper inzicht te verkrijgen in de wijze waarop de leerlingen reageerden op en omsprongen met de diverse componenten van het experimenteel programma. Daartoe werden alle lessen in de E-klas op video opgenomen. Aan de hand van deze video-opnamen en van al het schriftelijk leerlingmateriaal werd er achteraf een gedetailleerde reconstructie gemaakt van het concrete verloop van de lessen.

3 Hypothesen en bijkomende onderzoeksvragen

3.1 Hypothesen

Ten eerste, in de lijn van de resultaten van het onderzoek van Verschaffel e.a. (1994) werd er in de drie groepen een erg gering aantal 'realistische reacties' (RRs) op de P-items uit de voortoets verwacht. Ook voorspelden we dat

het aantal RRs van de vijfdeklassers uit de E-groep niet significant zou verschillen van dat van de zesdeklassers uit de twee controlegroepen, omdat deze laatsten tot dan toe niet méér instructie gekregen hadden in realistisch modelleren en interpreteren van contextgebonden toepassingen van de vier basisbewerkingen dan de leerlingen uit de E-groep.

Ten tweede verwachtten we in de E-klas een betekenisvolle toename van het aantal RRs op de P-items tussen de voor- en de natoets. In de C2-klas werd er geen stijging van het aantal RRs tussen de voor- en de natoets voorspeld. In de C1-klas daarentegen viel er wel een zekere toename van het aantal RRs te verwachten (als gevolg van de korte introductie op de natoets), doch de voorspelling luidde dat deze toename niet significant zou zijn. De argumentatie hiervoor was dat louter waarschuwen voor niet-realistische antwoorden niet zou volstaan om de sterke dispositie van leerlingen tot realistisch modelleren en interpreteren op betekenisvolle wijze positief te beïnvloeden.

Een derde hypothese was dat de leerwinst in de E-klas niet beperkt zou blijven tot de opgaven die goed lijken op de oefeningen waarmee in de experimentele leergang was gewerkt, maar dat er ook transfer zou optreden van de verworven disposities en vaardigheden naar probleemsituaties die duidelijk afwijken van de situaties waarmee tijdens het experimentele programma gewerkt was. Daarom voorspelden we dat de toename van het aantal RRs in de E-klas tussen de voor- en de natoets niet alleen bij de vijf leeritems maar ook bij de vijf nabije-transferitems significant zou zijn.

Ten vierde werd uitgegaan van de hypothese dat de leerwinst van de E-groep méér zou zijn dan een tijdelijke gedragsverandering. Derhalve luidde de voorspelling dat er in de E-klas geen terugval zou zijn wat betreft het aantal RRs op de nabije-transferitems uit de retentietoets in vergelijking met de natoets en dat ook op de vijf verre-transferitems beduidend meer RRs zouden aangetroffen worden dan in de vroegere constaterende studies waarin diezelfde P-opgaven aan groepen niet-getrainde leerlingen werden voorgelegd.

De eerste drie voorspellingen werden getoetst met behulp van een $3 \times 2 \times 2$ variantie-analyse met 'klasgroep' (E versus C1 versus C2), 'afna-

memoment' (voortoets versus natoets) en 'opgavetype' (leeritems versus nabije-transferitems) als onafhankelijke variabelen, en het aantal RRs als afhankelijke variabele. Significante effecten werden nader onderzocht met behulp van de Tukey a posteriori toets (met een significantieniveau van 5%). Omdat de retentietoets alleen in de E-groep werd afgenomen en bovendien slechts voor de helft bestond uit parallel-opgaven van de P-items uit de voor- en de natoets, werden de resultaten op de retentietoets niet betrokken in de variantie-analyse.

3.2 Bijkomende onderzoeksvragen

Naast het toetsen van de bovenvermelde onderzoekshypothesen, werd door middel van een gedetailleerde analyse (1) van de leerlingenantwoorden op de diverse toetsen en (2) van het concrete verloop van de experimentele leergang in de E-klas ook nog gepoogd om een antwoord verkrijgen op de volgende vragen:

- Had het experimenteel programma een vergelijkbaar (positief) effect op de 'sterke', de 'middelmatige' en de 'zwakke' leerlingen uit de E-klas?
- Was het effect identiek voor de vijf soorten van modellerings- en interpretatiemoeilikheden die in het programma aan bod kwamen?
- Traden er bij de leerlingen uit de E-klas verschillen op in de manier waarop de P-items (en de S-items) voor en na de training concreet werden aangepakt en/of beantwoord?
- Gaven de opgaven uit het experimenteel programma effectief aanleiding tot de bedoelde discussies over en reflecties op realistisch modelleren en interpreteren?
- Hoe sterk diende de lesgever deze groeps- en klasdiscussies te sturen en hoe reageerden de leerlingen daarop?

4 Resultaten in verband met de vier hypothesen

Alvorens de resultaten in verband met de vier bovenvermelde hypothesen te vermelden, signaleren we dat een betrouwbaarheidsonderzoek van de gehanteerde toetsen voor de P-items uit de voortoets, de natoets en de retentietoets respectievelijk een consistentiecoëfficiënt (KR20) van .54, .79 en .73 ople-

verde. De lage consistentie-coëfficiënt voor de voortoets is o.m. te wijten aan het feit dat er op verscheidene P-items uit deze voortoets door geen enkele of slechts één enkele leerling 'op realistische wijze' werd gereageerd (zie verder).

Een eerste onderzoeksresultaat houdt verband met de prestaties van de leerlingen op de voortoets. Zoals voorspeld in de *eerste hypothese* en overeenkomstig de resultaten van het onderzoek van Verschaffel e.a. (1994), demonstreerden de leerlingen uit de E-klas en de twee controleklassen een zeer sterke neiging tot het uitsluiten van context-gebonden kennis en realistische overwegingen bij het oplossen van de vraagstukken uit de voortoets. Voor de drie klassen (E, C1 en C2) samen bedroeg het percentage RRs op de tien P-items uit deze voortoets slechts 15%.⁴ Het percentage RRs van de vijfdeklassers uit de E-klas lag weliswaar lager dan dat van de zesdeklassers uit C1 en C2 - de percentages bedroegen respectievelijk 7, 20 en 18% - doch Tukey a posteriori toetsen wezen uit dat deze drie percentages niet significant van elkaar verschilden.

Het toetsen van de *tweede hypothese* over het effect van het experimenteel programma geschiedde door middel van een variantie-analyse met '(klas)groep' en '(afname)-moment' als onafhankelijke variabelen en 'aantal RRs' als afhankelijke variabele. Tevens werd via een aantal Tukey a posteriori toetsen voor elk van de drie (klas)groepen afzonderlijk nagegaan of er een significant verschil was tussen het totaal aantal RRs op de P-items uit de voor- en de natoets. De variantie-analyse bracht een significant 'groep $\times$ moment' interactie-effect aan het licht, $F(2, 50) = 16.71$, $MSe = 18.4128$, $p < .0001$. Uit de Tukey a posteriori toetsen bleek dat dit 'groep $\times$ moment' interactie-effect te wijten was aan de veel grotere toename van het aantal RRs in de E-groep dan in de beide controlegroepen. In de E-groep was er een significante stijging van het percentage RRs van 7% naar 51%. In de twee controleklassen daarentegen werd slechts een stijging van 20% naar 34% (voor C1) en van 18% naar 23% (voor C2) geconstateerd (zie Tabel 2). Deze stijging was in beide gevallen niet significant. De eerder geringe en niet-significante vooruitgang van de eerste controlegroep - in vergelijking met de E-groep - vormt een dui-

delijke indicatie dat alleen een waarschuwing om bij het oplossen van (school)vraagstukken rekening te houden met reële, context-gebonden aspecten van het probleem, niet volstaat om van stereotiepe en niet-realistische probleemoplossers kritische en realistische denkers te maken. Verder signaleren we dat er in de E-klas geen significante correlatie was tussen het aantal RRs tijdens de voor- en de natoets. Gezien het zeer geringe aantal RRs in de E-klas tijdens de voortoets is dit niet verwonderlijk. In de twee controleklassen daarentegen was er wel een significante samenhang (op het 1% niveau) tussen de resultaten op de voor- en de natoets: de Pearson correlatiecoëfficiënt voor C1 en C2 bedroeg respectievelijk .70 en .68.

Tabel 2

Aantal en percentage realistische reacties (RRs) op de 10 P-items uit de voor- en de natoets in de E-, C1- en C2-groep

Groep	Realistische reacties			
	Voortoets		Natoets	
	N	%	N	%
E (n=18*)	13	7%	91	51%
C1 (n=18)	36	20%	61	34%
C2 (n=17)	32	18%	39	23%

* Eén van de leerlingen uit de E-klas kon wegens ziekte niet deelnemen aan de natoets; dit brengt het aantal leerlingen uit de E-klas op 18 in plaats van 19.

Met het oog op het toetsen van de *derde hypothese*, die betrekking had op de wendbaarheid van de bereikte leerresultaten, werd het effect van het experimenteel programma op de vijf leeritems en de vijf nabije-transferitems afzonderlijk geanalyseerd. De verwachting was dat de leerwinst van de leerlingen uit de E-groep niet beperkt zou blijven tot de leeritems, maar zich ook zou uitstrekken tot de nabije-transferitems. De variantie-analyse bracht geen '(klas)groep $\times$ (afname)moment $\times$ (opgaven)type' interactie-effect aan het licht. Zoals Tabel 3 laat zien, was de vooruitgang van de leerlingen uit de E-klas op de vijf leeritems weliswaar groter dan op de vijf nabije-transferopgaven (resp. van 9% naar 60% en van 6% naar 41%), doch uit de Tukey a posteriori toetsen bleek dat de toename ook in het tweede geval significant was (op het 5% niveau). Met andere woorden, de vooruitgang van de E-groep bleef niet beperkt tot P-items die opgebouwd

Tabel 3

Aantal en percentage realistische reacties (RRs) op de vijf leeritems en op de vijf nabije-transferitems uit de voor- en de natoets in de E-, C1- en C2-groep

Groep	Realistische reacties							
	Voortoets				Natoets			
	Leeritems N	%	Nabije transferitems N	%	Leeritems N	%	Nabije transferitems N	%
E (n=18*)	8	9%	5	6%	54	60%	37	41%
C1 (n=18)	18	20%	18	20%	36	40%	25	27%
C2 (n=17)	16	20%	16	20%	19	22%	20	23%

* Eén van de leerlingen uit de E-klas nam wegens ziekte niet deel aan de natoets en werd derhalve niet opgenomen in deze tabel.

waren rond een probleemsituatie die een duidelijke verwantschap vertoonde met bepaalde opgaven uit het experimenteel programma, maar strekte zich ook uit tot die items uit de natoets die te maken hadden met contexten die daar sterker van afweken. Hieruit blijkt alvast dat de leerwinst van de E-groep zeker niet kan worden afgedaan als een simpel en specifiek trainings-effect.

Aanvullende empirische steun voor deze interpretatie van de leerwinst van de leerlingen van de E-klas, wordt geleverd door de uitslagen op de retentietoets die een maand na het beëindigen van het programma van deze leerlingen afgenomen werd (*hypothese 4*). Zoals in paragraaf 2.4.3 beschreven, bestond deze retentietoets uit vijf P-items die isomorf waren met de vijf nabije-transferopgaven uit de natoets en daarnaast ook nog uit vijf niet-isomorfe verre-transferitems waarvan uit recent onderzoek eveneens gebleken was dat zij door traditioneel geschoolde 10 à 14-jarigen slechts zelden op realistische wijze beantwoord worden. De verwachting was dat er in de E-groep geen terugval zou zijn wat betreft het aantal RRs op de P-items uit de retentietoets in vergelijking met de natoets. Het percentage RRs voor de vijf nabije-transferitems uit de retentietoets bedroeg 40%; dit is evenveel als voor de vijf overeenkomstige nabije-transferitems uit de natoets (nl. 41%, zie Tabel 2). Maar bovendien werd ook voor de vijf verre-transferitems uit de retentietoets een vergelijkbaar percentage RRs vastgesteld, nl. 39%. Dit percentage is beduidend hoger dan werd vastgesteld in andere recente studies (zie o.m. Greer, 1993; Verschaffel e.a., 1994) op identieke of sterk verwante opgaven bij vergelijkbare groepen leerlingen die geen speciale instructie in realistisch modelleren en interpreteren ontvingen.⁵ De Pear-

son correlatiecoëfficiënt tussen de uitslag van de leerlingen uit de E-klas op de natoets als geheel en de retentietoets bedroeg .74 ($p < .01$).

5 Resultaten in verband met de bijkomende onderzoeksvragen

Zoals aangekondigd in paragraaf 3.2, vergeleken we ook het effect van het experimenteel programma bij leerlingen met een verschillend niveau van rekenvaardigheid én bij de verschillende soorten modellerings- en interpretatiemoeilijkheden. Verder werd ook nog nagegaan of de (realistische) antwoorden en aanpakwijzen van de leerlingen uit de E-klas tijdens de natoets en de retentietoets kwalitatief verschilden van die uit de voortoets en voerden we een gedetailleerde analyse door van het concrete verloop van de lessen in de experimentele klas.

Om enig idee te verkrijgen van het verband tussen het niveau van rekenvaardigheid en de behaalde leerwinst, berekenden we de percentages RRs voor de P-items uit de voortoets, de natoets en de retentietoets voor de 'sterke', de 'middelmatige' en de 'zwakke' leerlingen uit de E-klas afzonderlijk.⁶ Zoals uit Tabel 4 blijkt, boekten de leerlingen uit de 'sterke' groep veruit de grootste leerwinst.

Daarnaast berekenden we het percentage RRs van de leerlingen uit de E-klas voor de vijf verschillende soorten van P-items uit de voortoets, de natoets en de retentietoets afzonderlijk (zie Tabel 5). De resultaten suggereren dat het programma niet even effectief was voor alle soorten P-items. Zoals uit Tabel 5 blijkt, was de toename van het percentage RRs veruit het grootst bij de items die te maken hadden met

Tabel 4

Aantal (N) en percentage (%) realistische reacties (RRs) op de P-items van de voortoets, de natoets en de retentietoets voor de wiskundig 'sterke', 'middelmattige' en 'zwakke' leerlingen uit de experimentele klas

Niveau	Voortoets		Realistische reacties Natoets		Retentietoets	
	N*	%	N	%	N	%
Zwak (n=6)	4	7%	20	33%	14	23%
Middelmattig (n=6)**	4	7%	24	40%	20	33%
Sterk (n=6)	5	8%	47	78%	37	62%

* De voortoets, de natoets en de retentietoets bestonden allemaal uit 10 P-items.

** Wegens ziekte kon één van de zeven leerlingen uit de middelmattige groep niet deelnemen aan de natoets.

het adequaat interpreteren van de uitkomst van een niet-opgaande deling (zie item 1a en 1b uit Tabel 1).

De vergelijking van de preciese aard van de RRs die door de leerlingen uit de E-klas voor en na de training werden gegeven, bracht volgende verschillen aan het licht (voor details zie Verschaffel, 1995). Ten eerste kwam het tijdens de natoets en de retentietoets relatief gezien veel vaker voor dat een 'realistisch antwoord' (RA) op een P-item vergezeld was van een schriftelijke commentaar waarin expliciet gewezen werd op de modellerings- of interpretatiemoeilijkheid; deze realistische commentaren waren over het algemeen ook veel duidelijker geformuleerd dan de (schaarse) realistische opmerkingen die tijdens de voortoets werden gemaakt. Een tweede verschilpunt betrof de berekeningswijzen van de leerlingen. Tijdens de voortoets troffen we op de antwoordbladen van de leerlingen uit de E-klas uitsluitend sporen aan van het gebruik van procedures voor cijferrekenen of voor gestandaardiseerd hoofd-

rekenen. Tijdens de natoets en de retentietoets werden bij 1/3 van de leerlingen uit de E-klas duidelijke sporen gevonden van het gebruik van meer informele werkwijzen (zoals 'uitvoeren tellen' bij de toetsitems die betrekking hadden op het thema van de tweede LE; zie item 2a en 2b uit Tabel 1) en/of van tekeningen of schetsen van de probleemsituatie (met name bij de toetsitems die betrekking hadden op het thema van de vierde LE; zie item 4a en 4b uit Tabel 1). Ten derde kwam het tijdens de natoets ook enkele keren voor dat een leerling uit de E-klas een (ongepaste of ronduit foutieve) 'realistische reactie' gaf op één van de S-items. Zo beantwoordden twee 'zwakke' leerlingen uit de E-klas het volgende S-item uit de natoets "Chris verzamelt strips. Hij heeft 19 boeken van Buck Danny, 27 boeken van Suske en Wiske en 15 strips van Blueberry. Hoeveel strips telt Chris' verzameling?" met "Dat kan je niet weten, want er kunnen dubbele strips tussen zitten". Wellicht baseerden deze leerlingen zich voor het geven van deze 'realistische reac-

Tabel 5

Aantal (N) en percentage (%) realistische reacties (RRs) van de leerlingen uit de experimentele klas (n = 18) op de vijf verschillende soorten van P-items uit de voortoets, de natoets en de retentietoets

Opgavetype	Voortoets		Realistische reacties Natoets		Retentietoets*	
	N	%	N	%	N	%
LE 1**	8	22%	30	83%	16	89%
LE 2	0	0%	20	56%	6	33%
LE 3	5	14%	13	36%	4	22%
LE 4	0	0%	16	44%	5	28%
LE 5	0	0%	12	33%	5	28%

* Alleen de vijf nabije-transferitems uit de retentietoets zijn in deze tabel opgenomen. De vijf verre-transferitems zijn niet opgenomen, omdat de onderliggende modellerings- of interpretatiemoeilijkheid niet (duidelijk) onder te brengen is in één van de vijf topics die in het experimenteel programma aan bod kwamen.

** De topics van de vijf leereenheden zijn: interpreteren van de uitkomst van een niet-opgaande deling (LE 1), modelleren van probleemsituaties rond sets met gemeenschappelijke elementen (LE 2), modelleren van probleemsituaties waarin het twijfelachtig is of een rekenkundige bewerking met de gegeven getallen de exacte uitkomst oplevert (LE 3), rekening houden met relevante contextuele aspecten die niet expliciet in de opgave vermeld staan (LE 4), en discrimineren tussen probleemsituaties die wel en niet gemodelleerd kunnen worden via lineaire proportionaliteit (LE 5).

tie' op de oppervlakkige gelijkenis tussen dit 'probleemloze' S-item en het 'problematisch' uitgangsprobleem van LE 2 van het experimenteel programma, dat eveneens over een strip-verzameling ging (zie paragraaf 2.3.1).

De analyse van de videobeelden van de lessen in de E-klas en van al de leerlingannotaties hielp ons niet enkel om sommige kwantitatieve onderzoeksresultaten te verklaren, maar leverde tevens belangrijke aanknopingspunten op voor het verbeteren van de diverse aspecten van het experimenteel programma. We beperken ons tot twee belangrijke vaststellingen (voor meer details zie Verschaffel, 1995). Ten eerste stelden we vast dat sommige P-items uit het experimentele programma aanleiding gaven tot overloze discussies tussen leerlingen of groepen van leerlingen die dit item wél en niet realistisch geïnterpreteerd en opgelost hadden. Vooral (groepen van) leerlingen die aanvankelijk een niet-realistisch antwoord hadden gegeven, gingen tijdens het klasgesprek geregeld ijverig op zoek naar (extreme, vergezochte) condities waaronder hun oorspronkelijk antwoord - dat op routine-matige wijze tot stand gekomen was - toch nog zou kunnen opgaan. Zo trachtten de leerlingen uit de groep die de tweede vraag van het uitgangsprobleem van LE1 met '38' in plaats van '37 kisten' geantwoord had, zich nog te verdedigen door op te werpen dat "vier zeer sterke soldaten samen misschien nog wel één kist konden dragen" en opperde een andere dat er "misschien wel één lichtere kist bij was", en kostte het de andere leerlingen (en de lesgever) heel wat moeite om deze leerlingen ervan te overtuigen dat dergelijke assumpties strijdig waren met de informatie die in de opgavetekst vermeld staat. Het regelmatig voorkomen van dergelijke overloze discussies is ons inziens een rechtstreeks gevolg van de aard van de problemen waarmee in het experimenteel programma is gewerkt: in vele gevallen waren deze problemen bewust gekozen en geformuleerd om de traditioneel geschoolde leerling 'op het verkeerde been te zetten'. Blijkbaar probeerden sommige leerlingen, nadat zij beseften dat ze 'in de val gelopen waren', zich achteraf nog te redden door op zoek te gaan naar contextuele condities waaronder hun oorspronkelijk (niet-realistisch antwoord) kon worden gehandhaafd. Ten tweede, bleek uit de analyse van de

videobeelden van de lessen in de E-klas en van het schriftelijk leerlingenmateriaal dat niet alle leerlingen even actief betrokken waren bij de lessen. Zowel tijdens het groepswork als bij de klasdiscussies namen enkele (knappere) leerlingen sterk het voortouw.

6 Discussie en perspectieven

Het doel van onderhavige studie bestond erin een antwoord te geven op de vraag of het mogelijk is om betekenisvolle positieve leereffecten te bereiken op het vlak van realistisch modelleren en interpreteren van wiskundige toepassingsproblemen door basisschoolleerlingen onder te dompelen in een nieuwe, krachtige leeromgeving waarin vraagstukken opgevat worden als oefeningen in realistisch modelleren en interpreteren, eerder dan als oefenstof voor het indrillen van stereotiepe, formele oplossingsmethoden. Deze studie geeft daar een positief antwoord op. Immers, de experimentele leergang leidde tot een betekenisvolle ontwikkeling van de dispositie en de vaardigheid van leerlingen uit de E-groep om bij het oplossen van vraagstukken rekening te houden met de reële context waarin het rekenprobleem is ingebed. De eerder geringe en niet-significante vooruitgang van de leerlingen uit de eerste controlegroep kan als een duidelijke indicatie worden beschouwd dat de gekende neiging van leerlingen om vraagstukken op niet-realistische wijze aan te pakken zeer weerbarstig is en derhalve niet via een beperkte en eenvoudige instructiemaatregel om te buigen is in een dispositie tot realistisch modelleren en interpreteren (zie ook Verschaffel e.a., ter perse).

Bij deze positieve onderzoeksresultaten moet echter een aantal kritische kanttekeningen worden gemaakt.

Ten eerste worden de gerapporteerde positieve resultaten van onderhavige studie enigszins gehypothekeerd door een aantal methodologische tekorten van deze studie. Zo is er de beperkte omvang van de experimentele groep en van de beide controlegroepen. Ook het onderzoeksinstrumentarium dat gehanteerd werd om de evolutie in de dispositie en de vaardigheid tot realistisch modelleren te evalueren, is voor verbreding vatbaar. We denken daarbij onder andere aan een verdere uitbreiding en di-

versificatie van de P-items uit de schriftelijke toets, maar ook aan het inschakelen van aanvullende instrumenten (zoals een individuele toets, een vragenlijst...) waarmee gepeild kan worden naar eventuele ontwikkelingen in de metacognitieve opvattingen en strategieën van de leerlingen met betrekking tot (het leren oplossen van) rekenvraagstukken.

Ten tweede beklemtonen we dat de P-items uit de natoets en de retentietoets door de leerlingen uit de E-klas alles bij elkaar genomen nog steeds vaak op een niet-realistische wijze beantwoord werden (zie Tabel 2 en 3), en dat het positief effect van de training niet bij alle leerlingen en niet bij alle soorten modellerings- en interpretatieproblemen even duidelijk was (zie Tabel 4 en 5). Uit de analyse van de video-beelden van de lessen en van het schriftelijk leerlingmateriaal is bovendien gebleken dat dit wellicht niet louter op rekening te schrijven is van de beperkte omvang van de training (nl. slechts vijf leereenheden van 2 à 2,5 uur), maar dat er zich daarnaast ook aanpassingen opdringen zowel in de aard van de opgaven waarmee gewerkt wordt (zoals: minder zgn. 'fop'-opgaven en meer échte, realistische problemen) als in de organisatie van het onderwijs (zoals: betere garanties inbouwen dat *alle* leerlingen ervan profiteren).

Tenslotte, hoewel het experimenteel programma in nauwe samenwerking met de leerkracht tot stand kwam en laatstgenoemde zelfs enkele kleinere onderdelen ervan voor z'n rekening nam (zie Verschaffel, 1995), blijft het zo dat deze leergang relatief los stond van het 'gewone' reken/wiskundeprogramma. We beklemtonen dat wij zeker geen voorstander zijn van de invoering in het reken/wiskunde-onderwijs van zo'n speciale lessenreeks waarin leerlingen op een intense en systematische manier getraind worden in het realistisch modelleren en interpreteren van contextgebonden wiskundige problemen, terwijl er in de overige rekenlessen enkel met stereotiepe en eenduidige opgaven wordt (verder)gewerkt. Integendeel, de ontwikkeling van de dispositie tot en vaardigheid in het realistisch modelleren en interpreteren van contextgebonden wiskundige toepassingsproblemen moet door het hele reken/wiskundeonderwijs gestimuleerd en ondersteund worden (zie ook Goffree, 1983; Gravemeijer, 1994; Greer, 1993; Treffers & De

Moor, 1990). In toekomstig onderzoeks- en ontwikkelingswerk kan niet voorbijgegaan worden aan de vraag hoe dit best concreet kan worden gerealiseerd.

Noten

- 1 Het woord tussen haakjes in '(remediërings)-programma' wijst op het feit dat dit programma zich speciaal richt op leerlingen die reeds meerdere jaren ervaring hebben opgedaan met traditioneel vraagstukkenonderwijs waarin weinig speciale aandacht uitgaat naar het leren realistisch modelleren en interpreteren.
- 2 Om praktische redenen was het niet mogelijk om de retentietoets ook in de beide controleklassen af te nemen.
- 3 Bij het categoriseren van reacties als RR of NR werden puur rekentechnische fouten buiten beschouwing gelaten. Wanneer een leerling bij het beantwoorden van een P-item realistisch tewerkgegaan was of bij een verwacht, niet-realistisch antwoord een realistische opmerking had geplaatst maar in zijn rekenwerk één of andere technische fout had gemaakt, werd zijn reactie op dit item dus gescoord als RR.
- 4 Dit is ongeveer evenveel als in het onderzoek van Verschaffel e.a. (1994), waarin een gelijkaardige toets bestaande uit tien P-items en tien S-items aan 75 vijfdeklassers uit drie verschillende scholen aangeboden werd, die in het totaal 17% RRs op de P-items uitlokte (zie paragraaf 1).
- 5 In de constaterende studie bij 75 vijfdeklassers van Verschaffel e.a. (1994) kwamen P-item 6, 7 en 10 uit de retentietoets eveneens voor. In deze studie werden deze opgaven respectievelijk door slechts 3%, 17% en 3% van de leerlingen op realistische wijze beantwoord.
- 6 Voor de plaatsing van de leerlingen in één van de drie vaardigheidsniveaus baseerden we ons op het oordeel van de leerkracht dat eveneens gebruikt was bij de vorming van de heterogene groepen voor de groepstaken.

Literatuur

- Cobb, P. (1989). Experiential, cognitive and anthropological perspectives in mathematics education. *For the Learning of Mathematics*, 9(2), 32-42.

- Corte, E. De, & Verschaffel, L. (1985). Beginning first graders' initial representation of arithmetic word problems. *Journal of Mathematical Behavior*, 4, 3-21.
- Corte, E. De, & Verschaffel, L. (1989). Teaching word problems in the primary school. What research has to say to the teacher. In B. Greer & G. Mulhern (Eds.), *New developments in teaching mathematics* (pp. 85-106). London: Routledge.
- Davis, R. B. (1989). The culture of mathematics and the culture of schools. *Journal of Mathematical Behavior*, 8, 143-160.
- Goffree, F. (1983). *Wiskunde en didactiek. Deel 2*. Groningen: Wolters-Noordhoff.
- Gravemeijer, K. (1994). *Developing realistic mathematics education*. Utrecht: CD-béta Press.
- Greer, B. (1993). The modeling perspective on word problems. *Journal of Mathematical Behavior*, 12, 239-250.
- Kilpatrick, J. (1987). Problem formulating: Where do good problems come from? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 123-147). Hillsdale, NJ: Erlbaum.
- Nelissen, J. (1987). *Kinderen leren wiskunde. Een studie over metacognitie en reflectie in het basisonderwijs*. Gorinchem: De Ruiter.
- Nesher, P. (1980). The stereotyped nature of school word problems. *For the Learning of Mathematics*, 1, 41-48.
- Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. *Instructional Science*, 17, 309-338.
- Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 311-343). Hillsdale, NJ: Erlbaum.
- Silver, E. A., Shapiro, L. J., & Deutsch, A. (1993). Sense making and the solution of division problems involving remainders: An examination of middle school students' solution processes and their interpretations of solutions. *Journal for Research in Mathematics Education*, 24, 117-135.
- Treffers, A., & Moor, E. de (1990). *Proeve van een nationaal programma voor het reken-wiskundeonderwijs op de basisschool. Deel 2. Basisvaardigheden en cijferen*. Tilburg: Zwijssen.
- Verschaffel, L. (1995). *Leren realistisch modelleren en interpreteren van vraagstukken. Verslag van een onderwijsonderzoek bij leerlingen van de bovenbouw van de basisschool*. (Intern rapport). Leuven: Centrum voor Instructiepsychologie en -Technologie, K.U.Leuven.
- Verschaffel, L., Corte, E. De, & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, 4, 273-295.
- Verschaffel, L., Corte, E. De, & Lasure, S. (in press). Children's conceptions about the role of real-world knowledge in mathematical modeling of school word problems. In W. Schnotz, S. Vosniadou & M. Carretero (Eds.), *Conceptual change*. Hillsdale, NJ: Erlbaum.
- Verschaffel, L., & Corte, E. De (in press). Word problems. A vehicle for authentic mathematical understanding and problem solving in the primary school? In P. Bryant & T. Nunes (Eds.), *How do children learn mathematics?* Hillsdale, NJ: Erlbaum.
- Verschaffel, L., & Gravemeijer, K. (1990). Contextrijk reken/wiskunde-onderwijs. In *Gids Basisonderwijs* (Curr. 7420, 1-32). Brussel: Samsom.

Manuscript aanvaard 23-11-1995

Auteurs

L. Verschaffel is Onderzoeksleider bij het Nationaal Fonds voor Wetenschappelijk Onderzoek, België en Deeltijds Hoogleraar aan de Faculteit der Psychologie en Pedagogische Wetenschappen van de K.U. Leuven.

E. De Corte is Gewoon Hoogleraar aan de Faculteit der Psychologie en Pedagogische Wetenschappen van de K.U. Leuven.

Correspondentie-adres: Centrum voor Instructiepsychologie en -Technologie, K.U. Leuven, Vesaliusstraat 2, B-3000 Leuven, België, tel: 32/16/326258, fax: 32/16/326274; e-mail: lieven.verschaffel@ped.kuleuven.ac.be

Teaching realistic mathematical modeling. An exploratory teaching experiment with children of the upper grades of the elementary school.

L. Verschaffel & E. De Corte. Pedagogische Studiën, 1996, 73, 322-337.

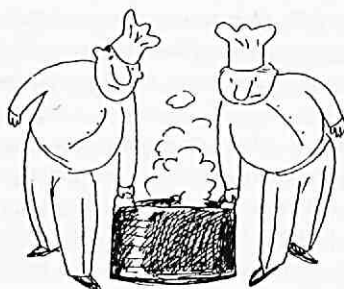
Recent research has convincingly documented elementary school children's tendency to neglect commonsense knowledge and realistic considerations during mathematical modeling of word problems in school arithmetic. The present paper describes the design and the results of a teaching experiment in which a group of 10-11-year olds followed a program wherein word problems are conceived as exercises in mathematical modeling, focussing on the assumptions and appropriateness of the model underlying any proposed solution. The program consists of five learning units of about 2.5 hours each, spread over a period of about 2.5 weeks. During the experiment, pupils from two control classes followed the regular mathematics curriculum. The three groups were given the same pretest and posttest, which contained word problems that could not be modeled in a straightforward and unproblematic way if one seriously takes into account the realities of the context called up by the problem statement. Besides specific training effects, transfer and retention effects were also investigated. The results of the experimental and control groups on the learning, transfer and retention tests provided convincing evidence for the effectiveness of the experimental program for developing in pupils a disposition towards realistic mathematical modeling.

LE 1 Uitgangsprobleem: Uit het soldatenleven
300 soldaten moeten per jeep naar het oefenterrein. Hoeveel jeeps zijn er nodig, als je weet dat er maximum 8 soldaten in één jeep kunnen plaatsnemen?



Als deze soldaten op het oefenterrein aankomen, worden ze naar een loods gebracht. In deze loods staan er een heleboel zware metalen kisten die naar een andere loods moeten worden gedragen. Om één zo'n kist te kunnen dragen, zijn er 8 man nodig. Hoeveel kisten kunnen er door deze 300 soldaten in één keer van de ene naar de andere loods gedragen worden?

Wanneer de soldaten terug in de kazerne zijn, hebben ze allemaal honger als een paard. De kok heeft voor hen 300 liter hutspot klaargemaakt. Hij heeft daarvoor 8 grote ketels gebruikt, die precies evenveel hutspot bevatten. Hoeveel liter hutspot zit er in één ketel?



Na het avondeten moeten de 300 soldaten nog marcheren. Ze moeten zich daarvoor opstellen in rijen van 8. Hoeveel soldaten blijven er over nadat er zoveel mogelijk rijen van 8 gevormd zijn?



LE 1 Uitgangsprobleem: Uit het soldatenleven
(antwoordenblad)

- 1 *Hoeveel jeeps zijn er nodig?*

Bewerking:

Antwoord:

- 2 *Hoeveel kisten kunnen er in één keer naar de andere loods gedragen worden?*

Bewerking:

Antwoord:

- 3 *Hoeveel liter hutspot zit er in één ketel?*

Bewerking:

Antwoord:

- 4 *Hoeveel soldaten blijven er over nadat er zoveel mogelijk rijen van 8 soldaten gevormd zijn?*

Bewerking:

Antwoord:

Bezinning

Is er jullie iets speciaals opgevallen bij het beantwoorden van deze vier vragen?

LE 1: Toepassingen

- 1 Op een familiefeest krijgen Paul, Sven, Kurt, Bram, Sarah en Charlotte van hun opa een doos waarin 75 ballonnen zitten. Zij verdelen die ballonnen eerlijk onder elkaar. Hoeveel ballonnen krijgt ieder?

Bewerking:

Antwoord:

- 2 Op de ouderavond van de school komen 81 ouders. Aan één tafel kunnen maximum 6 ouders zitten. Hoeveel tafels moeten er geplaatst worden?

Bewerking:

Antwoord:

- 3 Er worden 250 eieren in dozen van 12 ingepakt. Hoeveel volle dozen kunnen er worden gemaakt en hoeveel eieren blijven er over?

Bewerking:

Antwoord:

- 4 Vader heeft op de markt een kist appels gekocht voor 300 frank. Wanneer hij thuiskomt, weegt hij de appels. De weegschaal wijst 24 kg. aan. "Dat is dan ... frank per kg", zegt vader.

Bewerking:

Antwoord:

LE 1: Individuele opdracht

Maak zelf drie vraagstukken rond eenzelfde bewerking, nl. $100 : 8 =$.

Zorg ervoor dat het antwoord op het eerste vraagstuk '12' is, het antwoord op het tweede vraagstuk '13' is, en het antwoord op het derde vraagstuk '12,5'.

Veel succes!

Vraagstuk 1 (antwoord: 12)

Vraagstuk 2 (antwoord: 13)

Vraagstuk 3 (antwoord: 12,5)

Opleiding van ouders en de opvoeding van hun kinderen

Een onderzoek in de Kohn-traditie*

F. W. P. van der Slik, J. R. M. Gerris en A. J. A. Felling

Samenvatting

In dit artikel wordt een uitgebreid theoretisch model getest voor wat betreft de sociaal-structurele condities van ouders, hun opvoedingswaarden en hun opvoedingsgedragingen. Daarvoor is gebruik gemaakt van een representatieve steekproef van Nederlandse gezinnen ($N = 788$) met in ieder geval één kind tussen 9 en 16 jaar oud. Lisrel-analyses met een 'multi-group' benadering wezen uit dat de empirische modellen voor vaders en moeders niet van elkaar verschillen. Een tweede belangrijke uitkomst is dat uitsluitend de opleiding van ouders sterk significante effecten heeft op de opvoedingswaarden 'conformiteit aan externe autoriteiten' (Kohn) en op 'gehoorzaamheid aan de ouderlijke regels' (Gerris). Opleiding heeft echter geen effect op de derde in het onderzoek betrokken opvoedingswaarde: 'persoonlijke autonomie'. Bovendien bleek dat ouders die conformiteit aan externe autoriteiten een belangrijke opvoedingswaarde vinden, in mindere mate zelfstandigheid bij hun kinderen stimuleren, terwijl ouders die gericht zijn op gehoorzaamheid aan ouderlijke regels vaker hun kinderen straffen. Ouders die persoonlijke autonomie als opvoedingswaarde hanteren, tonen affectie naar hun kinderen, stimuleren hun zelfstandigheid en straffen minder. Verschillende verklaringen voor het massieve effect van de opleiding van de ouders worden besproken.

1 Introductie

De wijze waarop ouders hun kinderen opvoeden is van centraal belang voor de ontwikke-

lingskansen van hun kinderen. Zo zijn kinderen van warme en ondersteunende ouders niet alleen in sociaal en cognitief opzicht competent, maar ook emotioneel stabiel dan kinderen van koele en afstandelijke ouders (Belsky, 1981; Belsky, Robins & Gamble, 1984; Rollins & Thomas, 1979). Hoe ouders met hun kinderen omgaan blijkt tevens grote invloed te hebben op de mogelijkheden van kinderen om, wanneer zij eenmaal zelf volwassen worden, hechte relaties met anderen aan te gaan (cf. Kirkpatrick & Shaver, 1990). En tenslotte komt uit Amerikaans onderzoek naar voren dat de waarden, die kinderen van hun ouders meekrijgen, in belangrijke mate medebepalen tot welke sociale klasse de kinderen gaan behoren. Kinderen uit lagere sociale milieus worden voornamelijk opgevoed met waarden als gehoorzaamheid en conformiteit. Deze waarden zouden een rem zijn op een culturele integratie met hogere sociale strata, omdat daar juist waarden als autonomie en zelfbeschikking hoog in het vaandel staan (cf. McKinley, 1964; Kohn, 1989 [1969]).

Nu is bekend dat ouderlijke opvoedingswaarden en -gedragingen niet uitsluitend beïnvloed worden door persoonlijkheidskenmerken van de ouders (Belsky, 1984), door gezinssituationele kenmerken (Janssens & Gerris, 1988; Gerris, Vermulst & Siebenheller, 1990; Gerris, Vermulst & De Brock, 1991) en door kindkenmerken zoals leeftijd en temperament van het kind (Belsky, 1984; Vermulst, Gerris, Franken & Janssens, 1986). Ook sociaal-structurele factoren, zoals de sociale klasse van de ouders en de aard van hun beroep, bepalen in belangrijke mate op welke wijze de opvoeding van de kinderen gestalte wordt gegeven (zie voor een overzicht van empirische verbanden tussen sociale klasse en opvoedingsgedragingen: Gecas, 1979). Hoewel verschillende auteurs een interpretatie hebben gegeven van het verband tussen sociale klasse en

* Dit onderzoek maakt deel uit van het project 'Cultuur of klasse. Een onderzoek naar antecedenten van ouderlijke opvoedingswaarden en -gedragingen' (NWO-nr: 510-81-614).

ouderlijk opvoedingsgedrag lijkt de theorie van Melvin Kohn tot nu toe de meest adequate te zijn (cf. Gecas, 1979). We stellen daarom in dit artikel zijn theorie uitgebreid aan de orde.

2 De klassieke theorie van Kohn

Uitgangspunt van Kohn (1963, 1989) is dat ouderlijke waarden en opvattingen voortkomen uit ervaringen die mensen in het leven van alledag opdoen. Personen uit verschillende sociale klassen zouden te maken krijgen met uiteenlopende levens- en werkcondities. Het is om die reden dat leden van diverse sociale milieus de sociale werkelijkheid op verschillende wijzen zien en interpreteren, uiteenlopende toekomstverwachtingen koesteren en onderscheiden visies ontwikkelen omtrent wat waardevol en wenselijk is. Voor wat betreft de opvoeding kan men daarbij volgens Kohn onderscheid maken tussen twee fundamentele en elkaar uitsluitende waardenoriëntaties, te weten: zelfbepaling en conformiteit. Zelfbepaling of autonomie zou een typische hogere klasse waarde representeren, terwijl conformiteit of gehoorzaamheid een waarde is, die met name onderschreven zou worden door leden van lagere sociale strata.

De mate waarin ouders gericht zijn op deze opvoedingswaarden wordt volgens Kohn op een beslissende manier beïnvloed door het beroep en de daaruit voortvloeiende werkomstandigheden en werkervaringen. Zo zouden bijvoorbeeld beroepen van personen uit hogere sociale klassen meer gekenmerkt worden door het uitvoeren van zelfstandig genomen beslissingen, terwijl het werk dat door personen uit de lagere sociale strata wordt uitgevoerd meer routinematig van aard zijn en meer nadrukkelijk van bovenaf worden geleid. Kohns these is nu dat de ervaringen die mensen opdoen als gevolg van deze uiteenlopende werkomstandigheden ook tot verschillende opvattingen en waardenoriëntaties leiden. Dat zou niet alleen tot uitdrukking komen in meer of minder conservatieve opvattingen in de sociaal-culturele sfeer, maar ook terug te vinden zijn in de waarden die ten grondslag liggen aan de opvoeding van hun kinderen. Personen uit de lagere sociale strata zouden meer nadruk leggen op conformiteit aan externe regels, terwijl personen

uit de hogere echelons meer nadruk zouden leggen op zelfbeschikking en de ontwikkeling van interne gedragsstandaarden. En uiteindelijk zouden deze uiteenlopende opvoedingswaarden tot verschillende opvoedingsgedragingen leiden. Het is belangrijk op te merken dat deze uiteenlopende opvoedingswaarden functioneel zijn in de betreffende sociale en economische context. Er mag van uitgegaan worden dat alle ouders hun kinderen een zodanige opvoeding proberen te geven dat zij, wanneer ze eenmaal zelf volwassen zijn geworden, het goed zo niet beter zullen hebben dan zichzelf. Omdat ouders uit lagere sociale klassen op grond van hun eigen ervaringen bemerkt hebben dat conformiteit daarvoor een vereiste is, zullen zij deze waarde in de opvoeding van hun kinderen gestalte proberen te geven. Ouders uit hogere sociale klassen daarentegen zijn juist tot de bevinding gekomen dat autonomie van groot belang is om in de wereld vooruit te komen, met als gevolg dat zij deze waarde in de opvoeding van hun kinderen benadrukken.

Op grond van zijn resultaten komt Kohn tot de conclusie dat ouders afkomstig uit verschillende sociale klassen niet zozeer verschillen in de *mate* waarin zij hun kinderen straffen in overtredingssituaties, maar dat zij voornamelijk verschillen in de *wijze waarop* en de *omstandigheden waarin* ze hun kinderen tot de orde roepen. Ouders afkomstig uit de hogere sociale klassen zouden zich in de opvoeding vooral laten leiden door de intenties die volgens hen aan het gedrag van kinderen ten grondslag liggen, terwijl ouders uit de lagere klassen meer oog zouden hebben voor de onmiddellijke consequenties van het gedrag van hun kinderen. Een direct gevolg hiervan is dat ouders uit de hogere klassen, juist omdat ze gericht zijn op de ontwikkeling van interne gedragsstandaarden bij hun kinderen, in hun gedrag daadwerkelijk laten blijken dat zij onafhankelijkheid en zelfverantwoordelijkheid waarderen. Ouders daarentegen uit de lagere klassen, met hun oriëntatie op conformering aan externe regels, leggen in hun opvoedingsgedrag meer nadruk op controle. Het straffen in overtredingssituaties zou daarmee voor kinderen uit de hogere klassen anders uitwerken dan voor kinderen uit de lagere sociale klassen. Ouders uit hogere klassen streven ernaar controle via norminternalisatie bin-

nen het kind te leggen en mondt daarmee uit in zelfbepaling; ouders uit lagere sociale klasse leggen controle *buiten* het kind wat leidt tot conformiteit aan externe regels.

Op grond van uitgebreid empirisch onderzoek, zowel door hemzelf als door anderen uitgevoerd (zie Kohn, 1989 voor een overzicht), concludeert Kohn dat zijn theorie in belangrijke mate bekrachtigd wordt. Hoewel de gevonden samenhangen niet erg hoog zijn (de correlaties zijn zelden hoger dan .30), wijzen ze vrijwel zonder uitzondering in de voorspelde richting. Sociale klasse blijkt de belangrijkste predictor te zijn voor ouderlijke opvoedingswaarden en de sterkte van deze verbanden wordt nauwelijks beïnvloed door factoren als: leeftijd, ras, etniciteit, urbanisatiegraad, religiositeit (i.c. kerkgenootschap en kerkbezoek) en nationaliteit. Hij ontwikkelt derhalve een, op het eerste gezicht, succesvolle theorie waarop naderhand nogal wat kritiek is uitgeoefend.

In deze bijdrage zullen enkele kritische bedenkingen ten aanzien van de theorie van Kohn aan de orde worden gesteld. Op de eerste plaats is tot op heden onduidelijk gebleven wat de rol van het opleidingsniveau in dit geheel is. Zo kwamen Wright en Wright (1976) en Meijnen (1977) tot de bevinding dat het opleidingsniveau een zelfstandige en niet te verwaarlozen bijdrage leverde aan de verklaring van ouderlijke opvoedingswaarden. Meer nog, de samenhang tussen sociale klasse en Kohns conformiteitsindex wordt tot vrijwel nul gereduceerd wanneer gecontroleerd wordt voor de invloed van opleiding (Wright & Wright, 1976, p. 535). Hoewel Kohn uiterst geprikkeld reageerde op deze uitkomsten (Kohn, 1976), zijn door hem later vergelijkbare resultaten gerapporteerd - overigens zonder ook maar met één woord te verwijzen naar de resultaten van Wright en Wright (Slomczynski, Miller & Kohn, 1981). Om de directe invloed van opleidingsniveau op ouderlijke waardenoriëntaties te kunnen verklaren, veronderstelt Kohn (1983) dat reeds in de opleiding waarden als zelfbepaling en conformiteit overgebracht worden aan de leerlingen. Nog afgezien van de vraag of deze hypothese houdbaar is, impliceert de constatering dat het effect van sociale klasse op ouderlijke opvoedingswaarden en

-praktijken vrijwel tot nul gereduceerd wordt wanneer gecontroleerd wordt voor opleiding, dat Kohns theorie op een fundamenteel punt aan revisie toe is.

Op de tweede plaats kunnen vragen gesteld worden omtrent het belang dat door Kohn wordt gehecht aan de arbeidscondities voor de verklaring van opvoedingswaarden. In zijn vroegere werk gaat hij uit van de assumptie dat niet sociale klasse of opleiding een verklaring geven voor ouderlijke opvoedingswaarden en -gedragingen, maar de arbeidscondities die volgens hem het gevolg zijn van sociale klasse en opleiding (Kohn, 1963; Gecas, 1979). Met andere woorden: sociale achtergrondkenmerken als opleiding en sociale klasse zouden voorafgaan aan werkomstandigheden die op hun beurt weer de opvoedingswaarden beïnvloeden. In latere publikaties wordt echter verondersteld dat er een wisselwerking bestaat tussen arbeidscondities en sociaal-structurele indicatoren (Slomczynski et al., 1981). De 'status attainment theory' (cf. Blau & Duncan, 1967; De Graaf, 1987) heeft als assumptie dat arbeidscondities evenals het vereiste opleidingsniveau in belangrijke mate iemands sociale status bepalen. Het zal duidelijk zijn dat de plaats van werkomstandigheden in een theoretisch model ingrijpende consequenties heeft voor de theoretische verklaring van opvoedingswaarden. Immers, wanneer arbeidscondities causaal vooraf zouden gaan aan sociale klasse, kunnen zij alleen dan een centrale rol spelen in de verklaring van opvoedingswaarden voor zover ze een directe invloed laten gelden op ouderlijke opvoedingswaarden, terwijl tegelijkertijd een effect van sociale klasse op deze waarden moet ontbreken.

Ten derde kan opgemerkt worden dat Kohns theorie een sterk paternalistisch karakter draagt of op zijn minst erg eenzijdig op de arbeidscondities van de kostwinner is gericht. Slechts weinige onderzoekers zijn nagegaan wat de zelfstandige invloed is van de sociaal-structurele indicatoren van de moeder op haar opvoedingswaarden en -gedragingen, ondanks het feit dat de opvoedingstaak vroeger - maar ook tegenwoordig nog - voor het merendeel wordt uitgevoerd door vrouwen. In verreweg de meeste onderzoeken wordt de sociale klassepositie of beroepsstatus van de echtgenote bepaald aan de hand van de positie of status van haar man.

Zo'n benaderingswijze stuit evenwel op een fundamenteel bezwaar. Indien arbeidscondities werkelijk van centraal belang zouden zijn voor de verklaring van opvoedingswaarden, dan kunnen die werkomstandigheden uitsluitend een rechtstreekse invloed laten gelden op de opvoedingswaarden en -gedragingen van de beroepsbeoefenaar in casu de vader, niet op de waarden en gedragingen van zijn echtgenote. Immers, haar werkomstandigheden worden bepaald door haar gezin, met als gevolg dat haar opvoedingswaarden en -gedragingen sterk kunnen verschillen van die van haar man. Merkwaardigerwijs heeft voor zover wij weten geen enkele onderzoeker aan deze belangrijke consequentie van Kohns theorie aandacht besteed.

Een centraal thema in Kohns theorie is voorts dat sociale klasse en de daaruit voortvloeiende werkomstandigheden als verklaring dienen voor de mate waarin ouders gericht zijn op conformiteit dan wel zelfbeschikking, en dat deze verschillende oriëntaties verklaren waarom ouders uiteenlopende opvoedingsgedragingen hanteren. De door Kohn ontwikkelde conformiteitsindex voor opvoedingswaarden is daarmee een van de belangrijkste grootheden in het model omdat deze een interpretatie geeft voor het verband tussen sociale klasse en werkomstandigheden enerzijds en opvoedingsgedragingen anderzijds. Verschillende onderzoekers hebben evenwel hun twijfels geuit omtrent de betrouwbaarheid en de validiteit van deze waardenindex (Jansen, 1983; Van Liere & Winkels, 1987). Gezien het grote belang van ouderlijke opvoedingswaarden in het verklaringsmodel, zou het daarom wenselijk zijn indien de opvoedingswaarden op een meer uitgebreide wijze zouden kunnen worden geïndiceerd. Daardoor kan de betrouwbaarheid en de validiteit van de door Kohn ontwikkelde waardenindex binnen het verklaringsmodel meer uitgebreid worden onderzocht.

Vergelijkbare opmerkingen kunnen tenslotte gemaakt worden ten aanzien van ouderlijk opvoedingsgedrag. Daar waar voor opvoedingswaarden door Kohn nog een index is gebruikt, is dat bij opvoedingsgedrag niet het geval. Op grond van open interviews, waarin gedetailleerd is gevraagd naar de mate van overtredingen, naar het karakter van deze overtredingen en naar de reacties van ouders op dit

gedrag van hun kinderen, is door Kohn een classificatieschema opgesteld waarmee opvoedingsgedragingen van ouders uit verschillende klassen met elkaar zijn vergeleken. Nadeel van zo'n eclectische benadering is dat een aantal relevante onderscheidingen mogelijkerwijs niet of onvoldoende in het onderzoek wordt betrokken en dat bovendien gedragsmetingen aan de hand van afzonderlijke items onvoldoende betrouwbaar kunnen blijken te zijn (Fishbein & Ajzen, 1975).

3 Doel van het onderzoek

Doel van het onderzoek is tweeledig. Ten eerste zal nagegaan worden in hoeverre ook in Nederland een sociaal-structureel verklaringsmodel voor ouderlijk handelen teruggevonden kan worden. Er zijn op dit gebied reeds verschillende onderzoeken uitgevoerd (Gerris, Deković & Janssens, 1991a; Gerris, Deković & Janssens, 1991b; Janssens & Gerris, 1988; Vermulst et al., 1986) maar het betrof, anders dan in het hier uitgevoerde onderzoek, studies op grond van niet-representatieve gegevens. Ofschoon we ons zullen laten leiden door de inzichten van Kohn streven we geen replicatie van zijn onderzoek na. In navolging van Kohn veronderstellen we slechts dat ouderlijke opvoedingswaarden een volledige interpretatie kunnen geven voor het effect van sociaal-structurele kenmerken op ouderlijk opvoedingsgedrag. Hoewel onze studie in aanzet nogal bescheiden van opzet is, betekent dat nog niet dat we geen pretenties hebben. We veronderstellen namelijk enerzijds meer betrouwbare en meer valide maten voor ouderlijk opvoedingswaarden en -gedragingen te gebruiken en anderzijds empirische verklaringsmodellen voor zowel vaders als moeders te toetsen, waarbij expliciet zal worden nagegaan in hoeverre deze modellen onderling van elkaar verschillen.

Onze tweede intentie is de in het voorafgaande ter sprake gekomen kritische bemerkingen nader uit te werken. Deze kritiek kan vragenderwijs als volgt worden samengevat:

- * Welke sociaal-structurele factor is van groter belang voor de verklaring van ouderlijke opvoedingswaarden en -gedragingen: opleiding of sociale klasse?

- * Hebben arbeidsomstandigheden alleen een direct effect op ouderlijke opvoedingswaarden?
- * Wijkt het empirisch verklaringsmodel voor ouderlijk opvoedingshandelen van de vaders af van dat van de moeders en zo ja, is dat toe te schrijven aan uiteenlopende arbeidscondities?
- * Is de door Kohn ontwikkelde opvoedingswaarden-index een betrouwbaar meetinstrument?
- * Blijken alternatieve maten voor Kohns conformiteitsindex meer betrouwbare en meer valide maten op te leveren voor ouderlijk opvoedingshandelen?

4 Methode

4.1 Onderzoeksgroep

In de tweede helft van 1990 is onder een representatieve steekproef van Nederlandse gezinnen met in ieder geval een kind tussen 9 en 16 jaar, een grootschalig veldonderzoek gehouden met als doel inzicht te verkrijgen in het functioneren van zowel individuele ouders als het gezin in zijn geheel voor wat betreft het opvoeden van kinderen in het algemeen en hun kind tussen 9 en 16 jaar in het bijzonder. Van de 1829 benaderde gezinnen hebben uiteindelijk 788 gezinnen (response = 43%) aan het onderzoek deelgenomen. De uitval bleek niet selectief te zijn in die zin dat hij gelijkelijk verdeeld was naar regio en urbanisatiegraad. Voor het vergaren van het eerste deel van de gegevens zijn de in het onderzoek betrokken gezinsleden (vader, moeder en kind) individueel benaderd. Na afronding van dit eerste deel kregen de ouders een aanvullende set van vragenlijsten, die zij, na individueel ingevuld te hebben, konden retourneren in een portvrije antwoord-enveloppe. Gelet op het waarschijnlijkheidskarakter waarmee de steekproef is samengesteld zijn de uitkomsten van deze studie generaliseerbaar naar de Nederlandse gezinnen met in ieder geval een kind tussen 9 en 16 jaar. Voor meer gedetailleerde informatie omtrent opzet van het onderzoek, steekproeftrekking, representativiteit alsmede dataverzamelingsprocedure, zie: Gerris et al. (1992, pp. 7-11).

4.2 Meetinstrumenten

Sociaal-structurele indicatoren

In de vragenlijst is uitgebreid aandacht besteed aan de sociaal-structurele lokatie van beide ouders. Naast vragen omtrent opleiding, gezinsinkomen en het aantal uren per week dat men contractueel arbeidsplichtig is, is ook informatie ingewonnen over huidig dan wel vroeger beroep en functie, over de mate waarin men leiding geeft en of men al dan niet zelfstandig werkzaam is. De gegevens omtrent beroep, functie, al dan niet zelfstandig werkzaam zijn en de mate van leiding geven zijn gebruikt om de EGP-klassenindeling (Erikson, Goldthorpe & Portocarero, 1979) te genereren. De door ons gebruikte indeling van *sociale klasse* heeft zes categorieën (tussen haakjes vermelden we de oorspronkelijke klassenindeling van Erikson et al., 1979):

- Klasse 1: Hoger leidinggevend personeel, academici, grote zelfstandigen (I)
- Klasse 2: Lager leidinggevend personeel, geschoolde hoofdarbeiders, hogere technici (II)
- Klasse 3: Routine hoofdarbeiders (III)
- Klasse 4: Kleine zelfstandigen met personeel (IVA), kleine zelfstandigen zonder personeel (IVB), zelfstandige agrariërs en vissers (IVC)
- Klasse 5: Geschoolde handarbeiders, supervisors handarbeid, lagere technici (VVI)
- Klasse 6: Ongeschoolde handarbeiders (VIIA), landarbeiders (VIIB)

De gegevens omtrent *opleiding*, *gezinsinkomen* en *arbeidsduur* worden als aanvullende informatie over iemands sociaal-structurele lokatie meegenomen. Met nadruk zij er op gewezen dat we dus zowel voor de vaders als de moeders gegevens ter beschikking hebben over sociale klasse, opleiding, contractuele arbeidsduur, gezinsinkomen en ook over de hierna volgende indicatoren voor werkomstandigheden, opvoedingswaarden en opvoedingsgedragingen. Opgemerkt dient te worden dat 422 van de 788 moeders op het tijdstip waarop het onderzoek uitgevoerd werd geen betaald werk buitenshuis verrichtten. Van die 422 moeders heeft 93% daarvoor wel een baan gehad. Voor de EGP-klassenindeling zijn zowel de gegevens van personen met huidig werk als personen met vroeger werk gebruikt. Om expliciet

na te gaan of het al dan niet verrichten van betaald werk van invloed is op de ouderlijke opvoedingswaarden en -gedragingen hebben we daarom een variabele in het onderzoek meegenomen die dit onderscheid tot uitdrukking brengt. Deze variabele is benoemd als *werkend*.

Indicatoren arbeidscondities

Omdat volgens Kohn arbeidscondities een beslissende rol zouden spelen voor de verklaring van opvoedingswaarden en -gedragingen wordt in dit onderzoek nadrukkelijk aandacht besteed aan de werkomstandigheden. Autonomie op het werk is volgens Kohn een cruciale predictor voor de mate waarin men in de opvoeding gericht is op zelfbepaling dan wel conformiteit (Kohn, Naoi, Schoenbach, Schooler & Slomczynski, 1990). McKinley (1964) kwam zelfs tot de bevinding dat autonomie op het werk en arbeidssatisfactie een grotere invloed hadden op opvoedingsgedrag dan sociale klasse.

Enerzijds zijn objectieve maten voor werkomstandigheden in het onderzoek betrokken, te weten: de mate van leiding geven en het al dan niet zelfstandig werkzaam zijn. Deze indicatoren, die dus vooral de autonomie van het werk indiceren, zijn reeds in de EGP-klassen-indeling verdisconteerd en worden daarom niet als zelfstandige variabelen in het onderzoek betrokken. Anderzijds zijn subjectieve maten voor werkomstandigheden ontwikkeld. Aan de ouders met betaald werk is een groot aantal vragen gesteld omtrent de mate van tevredenheid over collega's, werkomstandigheden, zelfstandigheid in het werk, promotiemogelijkheden en omtrent werkstress. Op grond van deze informatie zijn aan de hand van Principale Factoren Analyse (extractiecriteria: $\lambda \leq 1$ en 'kniktest') zowel voor de vaders als voor de moeders twee Likert-schalen (range: 1.00 - 7.00) geconstrueerd. Deze zijn als volgt benoemd:

Arbeidssatisfactie. Deze uit zes items bestaande schaal ($\alpha = .67$ voor vaders; $\alpha = .66$ voor moeders) indiceert de mate waarin men tevreden is met de werkomstandigheden in het algemeen en de zelfstandigheid van het werk in het bijzonder. Representatieve items zijn: 'Ik ben tevreden over de zelfstandigheid die ik heb in mijn werk', 'Ik ben tevreden over mijn mogelijkheden om vooruit te komen' en

'Ik ben ontevreden over de werkomstandigheden'.

Werkstress. Hoge scores op deze schaal (twee items; $\alpha = .67$ voor vaders; $\alpha = .79$ voor moeders) indiceren dat men thuis het werk moeilijk van zich af kan zetten en dat men vindt dat het werk te veel druk oplegt.

Indicatoren opvoedingswaarden

Kohns conformiteitsindex. Een achttal, aan Kohn ontleende opvoedingswaarden is aan de ouders voorgelegd met de vraag of zij die ieder apart wilden rangschikken naar belangrijkheid. De in hun ogen meest belangrijke kreeg de score 1, de daaropvolgende score 2. De op één na minst belangrijke kreeg score 4 en de minst belangrijke score 5. De vier niet gekozen opvoedingswaarden kregen een score 3 toegerekend. Het gaat hier om opvoedingswaarden als: rekening houden met anderen, verantwoordelijkheidsgevoel kennen, verstandig en helder oordelen, goede manieren hebben, best doen op school. De Kohn-index meet de mate waarin ouders het belangrijk vinden dat het kind zich conformeert aan maatschappelijke conventies.

In navolging van Kohn (1989, p. 57) zijn Principale Componenten Analyses over deze rangorde-data uitgevoerd. Daarbij zijn uitsluitend de gegevens gebruikt van die ouders die een volledig scorepatroon (éénmaal score 5, 4, 2 en 1 en viermaal score 3) lieten zien. Deze analyses resulteerden zowel over de gegevens van de vaders als van de moeders in een één-factoroplossing (extractie-criteria: $\lambda \leq 1$ en 'kniktest'). In tegenstelling tot Jansen (1983) en De Witte (1990) vonden we geen aanwijzingen voor een twee-dimensionele structuur en evenmin verschilden zoals bij Jansen (1983) de factoroplossingen van ouders met een dochter als onderzoekseenheid van die van ouders met een zoon¹. Op basis van deze resultaten zijn waardenindices geconstrueerd door gemiddelde scores te berekenen over de acht uitspraken. Daarbij zijn de vier conformiteits-items zodanig gehercodeerd dat hoge scores conformiteit tot uitdrukking brengen. De aldus verkregen Kohn-indices (range: 2.25 - 3.75; $\bar{x} = 3.00$) hebben een lage betrouwbaarheid ($\alpha = .48$ voor vaders; $\alpha = .50$ voor moeders)².

Gezien deze lage betrouwbaarheden leek het ons wenselijk om te zoeken naar meer betrouwbare en tegelijkertijd meer handzame al-

ternatieven voor de Kohn-index³. Daarom is naast de aan Kohn ontleende opvoedingswaarden aan de ouders nog een groot aantal andere vragen voorgelegd omtrent waarden die zij van belang kunnen achten voor de opvoeding van hun kinderen. Omdat reeds in de opzet van het onderzoek van Gerris et al. (1993) uitgegaan is van een multi-operationele benadering zijn expliciet vragen opgenomen die als alternatief zouden kunnen dienen voor Kohns conformiteitsindex. Deze in Likert-format gestelde vragen zijn gedeeltelijk nieuw ontwikkeld en gedeeltelijk ontleend aan eerder onderzoek (cf. Vermulst et al., 1986; Vermulst, Gerris & Siebenheller, 1987). De antwoorden op deze vragen zijn met behulp van factoranalyses (PFA) onderzocht. Op grond van deze analyses zijn Likert-schalen (range: 1.00 - 7.00) geconstrueerd waarvan er twee als alternatieve operationalisaties van Kohns opvoedingswaarden-index opgevat kunnen worden (cf. Gerris et al., 1993). Vergelijkbaar met Jansen (1983) hebben we twee niet negatief met elkaar samenhangende schalen ($r_{1,2} = .07$, n.s. voor moeders; $r_{1,2} = .14$, $p \leq .05$ voor vaders) gevonden die als volgt zijn benoemd:

Persoonlijke autonomie. Deze schaal (zeven items; $\alpha = .83$, $\bar{x} = 6.21$ voor vaders; $\alpha = .83$, $\bar{x} = 6.33$ voor moeders) geeft de mate aan waarin ouders belang hechten aan het bereiken van zowel persoonlijke autonomie als een evenwichtige levenshouding. Daarbij moeten denken aan opvoedingswaarden als: een eigen mening hebben, onafhankelijkheid, zelfstandig zijn, verantwoordelijkheid kunnen dragen en doorzettingsvermogen hebben.

Conformiteit. Het betreft hier een uit 11 items bestaande schaal ($\alpha = .82$, $\bar{x} = 4.32$ voor vaders; $\alpha = .80$, $\bar{x} = 4.35$ voor moeders) die uitdrukking geeft aan de mate waarin ouders het belangrijk vinden dat het kind zich aanpast aan de eisen van de directe sociale omgeving. In volgorde van afnemende belangrijkheid (cf. Gerris et al., 1993, p. 84, p. 176) zijn dat: de eisen, regels en normen van de ouders zelf, vanuit het gezin, en vanuit de samenleving. Typische opvoedingswaarden zijn: 'Mijn zoon/dochter moet leren zich ten opzichte van zijn/haar ouders te gedragen zoals het hoort', 'Het is belangrijk dat kinderen gehoorzamen', 'Mijn zoon/dochter moet leren dat de regels en gebruiken in het eigen gezin belangrijk zijn' en

'Ik vind dat mijn zoon/dochter moet leren om te gaan met anderen zoals dat in ons gezin hoort'. Deze conformiteitsschaal is in diverse voorafgaande studies met uiteenlopende steekproeven valide gebleken (De Brock, 1994; Siebenheller, 1990; Deković, 1991).

Indicatoren opvoedingsgedragingen

In hedendaags gezinspedagogisch onderzoek wordt voor wat betreft opvoedingsgedragingen onderscheid gemaakt tussen twee globale concepten, te weten: ondersteuning en controle (zie voor een overzicht: Gerris et al., 1992). Ondersteuning als opvoedingsgedrag blijkt in onderzoek gerepresenteerd te worden door een één-dimensioneel construct, dat door onderzoekers verschillend wordt benoemd: warmte, affectie, liefde, betrokkenheid, intimiteit; controle door uiteenlopende en meer-dimensionele constructen. Naast de wijze waarop controle wordt uitgeoefend (induceren van empathie en/of schuld en angst) wordt er in de vakliteratuur onderscheid gemaakt tussen de intensiteit van controle (streng vs. niet streng), het doel van controle (kind aanzetten tot eigen verantwoordelijkheid en/of conformiteit aan geldende regels) en de kwaliteit van controle (redeneren, straf en/of liefdesonthouding).

Voor de meting van opvoedingsgedrag sluiten we aan bij recent onderzoek waarin gewerkt wordt met de constructen 'autoritatieve controle', 'ondersteuning' en 'restrictieve controle' (Ten Haaf & Janssens, 1991; Ten Haaf, 1993). Autoritatieve controle verwijst daarbij naar uitleggen en aanmoedigen van onafhankelijk gedrag, ondersteuning naar affectie en kind-gerichte reacties van de ouders, terwijl restrictieve controle aan verbale en fysieke bedreiging, straf en afwijzing refereert. Ten Haaf (Ten Haaf, 1993; Ten Haaf, Janssens & Gerris, 1994) heeft de validiteit van de door haar gebruikte gedragsschalen onderzocht aan de hand van 'Multi-Trait Multi-Method-analyses', waarbij verschillende ouderlijke opvoedingsgedragingen met behulp van diverse methoden (observaties, interviews, vragenlijsten en ratingsschalen) zijn gemeten. Op grond van haar resultaten kan geconcludeerd worden dat er sprake is van convergente validiteit van de schalen die de hierboven beschreven opvoedingsgedragingen beogen te meten, zij het dat de validiteit van de maten voor autoritatieve

controle bij de vaders geringer is. De voor ons onderzoek relevante conclusie moet zijn dat met behulp van vragenlijsten gemeten ouderlijk opvoedingsgedrag niet uitsluitend verwijst naar gedragsintenties maar naar gerapporteerd daadwerkelijk gedrag (Ten Haaf, 1993). In ons onderzoek wordt dan ook gewerkt met de volgende gedragsschalen: autonomie, affectie-expressie en straf. 'Autonomie' refereert aan het construct 'autoritatieve controle', affectie-expressie aan 'ondersteuning' en de schaal 'straf' vormt een indicator voor 'restrictieve controle'.

Autonomie. Met de gedragsschaal 'autonomie' (zeven items; $\alpha = .72$, $\bar{x} = 4.65$ voor vaders; $\alpha = .72$, $\bar{x} = 4.64$ voor moeders) wordt de mate achterhaald waarin ouders aangeven in hun interacties met hun zoon/dochter onafhankelijkheid, zelfverantwoordelijkheid voor het handelen, denken en beslissingen aan te moedigen en te benadrukken. Representatieve items van deze schaal zijn: 'Ik laat mijn zoon/dochter zijn/haar problemen vaak zelf oplossen', 'Ik laat mijn zoon/dochter veel nieuwe dingen uitproberen, ook al is de afloop van deze dingen heel onzeker'.

Affectie-expressie. Deze uit negen items bestaande schaal indiceert de mate waarin ouders aangeven in hun opvoeding voelbare en waarneembare uitingen van positieve affectie en genegenheid te hanteren. De betrouwbaarheden van deze gedragsschaal zijn zeer bevredigend ($\alpha = .87$, $\bar{x} = 4.35$ voor vaders; $\alpha = .87$, $\bar{x} = 4.75$ voor moeders). Kenmerkende items zijn: 'Ik knuffel, kus en omarm mijn zoon/dochter vaak', 'Ik geef mijn zoon/dochter vaak een complimentje'.

Straf (vijf items; $\alpha = .79$, $\bar{x} = 2.94$ voor vaders; $\alpha = .78$, $\bar{x} = 2.88$ voor moeders). De mate waarin ouders aangeven in hun opvoeding gebruik te maken van verschillende strafmaatregelen. Het gaat hier om straffen als: naar de eigen kamer sturen, iets leuks verbieden, draai om de oren of een tik.

5 Resultaten

Data-analyse

De gegevens zullen op twee verschillende wijzen geanalyseerd worden. Op de eerste plaats presenteren we bivariate samenhangen tussen

sociaal-structurele indicatoren en werkomstandigheden enerzijds en opvoedingswaarden en opvoedingsgedragingen anderzijds. Op de tweede plaats zullen we causale analyses uitvoeren met behulp van LISREL 7 (Jöreskog & Sörbom, 1990). Het voordeel van de laatste analyse is dat we kunnen nagaan in hoeverre de aanspraken van Kohns theorema dat alleen ouderlijke opvoedingswaarden een interpretatie geven van de invloed van sociaal-structurele kenmerken en arbeidscondities op opvoedingsgedragingen, door de werkelijkheid worden ondersteund.

Bivariate samenhangen

Voordat we de resultaten van de causale analyses presenteren zullen we verslag doen van de enkelvoudige samenhangen tussen sociaal-structurele indicatoren en arbeidscondities enerzijds en opvoedingswaarden en opvoedingsgedragingen anderzijds.

Wat uit de onderste helft van Tabel 1 onmiddellijk duidelijk wordt is dat we, evenals Meijnen (1977), geen sterk significante ($\alpha \leq .01$) samenhangen tussen werkomstandigheden en de hier onderzochte opvoedingsaspecten hebben gevonden. Deze uitkomsten staan diametraal tegenover die van McKinley (1964), die op basis van zijn gegevens concludeert dat arbeidssatisfactie zelfs een sterkere invloed zou hebben op opvoedingsgedragingen dan sociale klasse.

Als we vervolgens onze aandacht richten op de samenhangen in de bovenste helft van Tabel 1 dan kan het volgende geconstateerd worden. Zowel gezinsinkomen, opleiding als het al of niet werkend-zijn tonen negatieve samenhangen met Kohns conformiteitsindex en met de alternatieve schaal voor conformiteit. Naarmate het gezinsinkomen hoger is en naarmate ouders een hogere opleiding hebben zijn ze in hun opvoeding minder op conformiteit gericht. Bovendien blijken werkende ouders minder conformistische opvoedingswaarden te hanteren dan niet-werkende ouders. Aanvullende analyses (SPSSx: ONEWAY) toonden aan dat de sterkte van de in Tabel 1 gepresenteerde associatiematen (η) van de nominale variabele sociale klasse met Kohns conformiteitsindex en de conformiteitsschaal voornamelijk toe te schrijven is aan het feit dat lagere (en hogere) leidinggevend (klasse 2 en 1) significant la-

Tabel 1

Samenhang tussen sociaal-structurele indicatoren en arbeidscondities enerzijds en opvoedingswaarden en -gedragingen anderzijds: (Pearson's r en η (cursief) voor samenhangen met nominale variabele: sociale klasse)

	WADERS/MOEDERS ^a	Opvoedingswaarden				Opvoedingsgedragingen		
		Kohns conformiteitsindex	Persoonlijke Autonomie	Conformiteit	Autonomie	Affectie-expressie	Straf	
Sociaal-structurele indicatoren	Sociale klasse	.26**	.20**	.13	.27**	.20**	.19**	.13
	Gezinsinkomen	-.27**	-.18**	-.03	-.22**	-.15**	.17**	.08
	Arbeidsduur	-.03	-.04	.01	.05	-.08	.01	.12*
	Werkend	-.17**	-.10*	-.01	-.12**	-.15**	-.02	.09*
Indicatoren arbeidscondities	Opleiding	-.31**	-.30**	-.07	-.35**	-.28**	.19**	.14**
	Arbeidsatisfactie	.05	.10	.06	.02	.12*	.01	.07
	Werkstress	-.09*	-.06	-.02	-.09*	-.08	.06	.11*
		(*: $\alpha \leq .05$; **: $\alpha \leq .01$ (tweezijdig getoetst) a: linkerwaarde voor de vaders; rechterwaarde voor de moeders						
		(Vaders: 527 $\leq N \leq$ 717) (Moeders: 270 $\leq N \leq$ 773)						

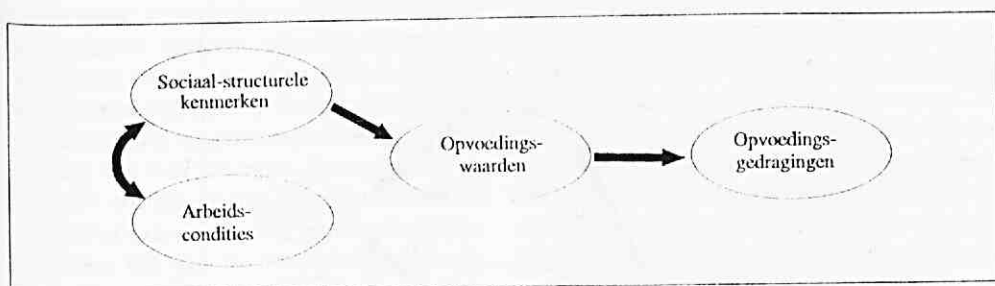
gere gemiddelde scores hebben op deze maten dan ongeschoolde (en geschoolde) handarbeiders (klasse 6 en 5). De scores van hoofdarbeiders (klasse 3) en van kleine zelfstandigen (klasse 4) lieten geen significante verschillen zien. Hoewel deze uitkomsten dus in lijn liggen van Kohns these moet onmiddellijk worden toegevoegd dat de verbanden met opleiding telkens sterker zijn dan die met sociale klasse. Dit komt duidelijk overeen met de uitkomsten van Wright en Wright (1976). Tevens is te zien dat de genoemde achtergrondkenmerken geen samenhang vertonen met de alternatieve schaal voor Persoonlijke Autonomie. Ook de contractuele arbeidsduur vertoont geen samenhang met de opvoedingswaarden.

Ten slotte kan uit deze kruisverbanden worden opgemaakt dat de mate waarin ouders aangeven hun kinderen te straffen in overtredings-situaties significant samenhangt met opleiding, sociale klasse en inkomen, waarbij de samenhangen met opleiding en sociale klasse het hoogst zijn. Hoe hoger de opleiding en hoe hoger het gezinsinkomen, hoe minder vaak ouders aangeven dat zij hun kinderen straffen. Lagere leidinggevendenden (klasse 2) bleken significant minder straf te geven dan ongeschoolde arbeiders (klasse 6), terwijl hoofdarbeiders (klasse 3) en kleine zelfstandigen (klasse 4) daar precies tussenin vallen.

De mate waarin ouders affectie tonen naar hun kinderen blijkt alleen bij de vaders een tamelijk zwakke associatie met sociale klasse te laten zien: lagere leidinggevendenden (klasse 2) tonen meer affectie naar hun kinderen dan ongeschoolde arbeiders (klasse 6). Ook de mate waarin ouders in hun gedrag gericht zijn op zelfverantwoordelijkheid bij hun kinderen blijkt bij de vaders samen te hangen met sociale klasse, gezinsinkomen en opleiding, terwijl bij de moeders alleen een significante samenhang met opleiding kan worden geconstateerd. Opmerkelijk is evenwel dat naarmate moeders meer uren per week arbeidsplichtig zijn dan wel een werkkrijs hebben zij in hun gedrag meer de autonomie van hun kind benadrukken. Bij de vaders zijn deze samenhangen afwezig.

Causale analyses. Werkwijze

Bij het uitvoeren van de LISREL-analyse zijn we uitgegaan van het in Figuur 1 weergegeven theoretische model.



Figuur 1. Theoretisch model voor verklaring van opvoedingswaarden en -gedragingen

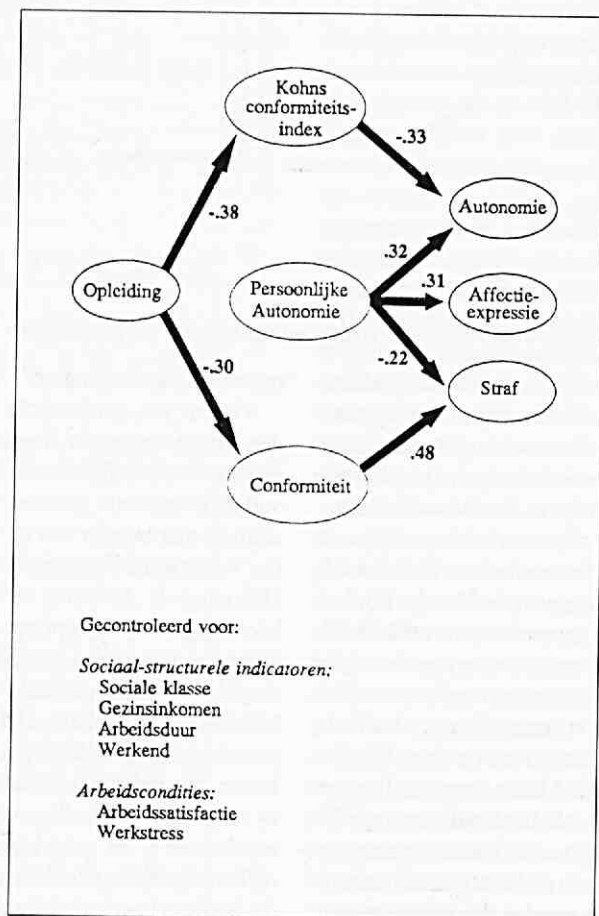
Omdat niet duidelijk is of arbeidscondities zoals arbeidssatisfactie en werkstress oorzaak of gevolg zijn van, dan wel in wisselwerking zijn met, sociale klasse, inkomen, opleiding en arbeidsduur hebben we in het theoretisch model een samenhang tussen arbeidscondities en sociaal-structurele kenmerken verondersteld. Ook tussen opleiding, sociale klasse en inkomen hebben we een samenhang verondersteld. Assumpties vanuit statusverwervingstheorieën omtrent causale samenhangen hebben we niet in ons model verwerkt omdat dit niet relevant is voor onze vraagstelling. We zijn namelijk uitsluitend geïnteresseerd in de vraag welke van deze variabelen een zelfstandige bijdrage leveren aan de verklaring van opvoedingswaarden.

Uit de analyses van de bivariate samenhangen kon opgemaakt worden dat arbeidscondities geen sterk significante samenhang vertonen met opvoedingswaarden. Omdat ook Meijnen (1977) geen effecten van werkomstandigheden op opvoedingswaarden heeft achterhaald, hebben we daarom in het theoretisch model geen effecten verondersteld van de arbeidscondities op ouderlijke opvoedingsoriëntaties. Tevens hebben we op grond van deze voorafgaande analyses en op basis van de uitkomsten van Wright en Wright (1976), Meijnen (1977), Williamson (1984) en Bosker, Meijnen en Van der Velden (1988) uitsluitend een invloed verondersteld van opleiding op opvoedingswaarden. En tenslotte zij opgemerkt dat we in ons theoretisch model geen directe invloed van sociaal-structurele factoren op opvoedingsgedragingen hebben toegelaten. We veronderstellen namelijk in navolging van Kohn, hoewel door hem niet expliciet getoetst, dat ouderlijke opvoedingswaarden een volledige interpretatie kunnen aandragen voor de invloed van sociaal-structurele kenmerken op

opvoedingsgedragingen.

Omdat we, conform de vraagstelling, wilden onderzoeken in hoeverre het empirisch model voor moeders verschilt van dat voor de vaders is expliciet gekozen voor een LISREL-analyse met behulp van de 'multi-sample' optie. Voor dat doel zijn met behulp van PRELIS (Jöreskog & Sörbom, 1990) covariantiematrices gegenereerd (option: pairwise deletion). Voor het testen van het hiervoor beschreven model is gebruik gemaakt van de 'Maximum Likelihood' procedure. Het gemiddelde aantal paarobservaties is daarbij als uitgangspunt gekozen. We hebben gewerkt met geobserveerde x- en y-variabelen die een één-op-één relatie met latente ζ - en η -variabelen hebben. Om zo dicht mogelijk aan te sluiten bij de realiteit zijn de betrouwbaarheden van de geobserveerde y-variabelen in het model verdisconteerd (Jöreskog & Sörbom, 1989, p. 136). De modificatie-indices zijn gebruikt om tot zo goed mogelijk passende modellen te komen. Bij het gebruik van deze modificatie-indices hebben we ons wel een restrictie opgelegd. We hebben alleen directe effecten van de sociaal-structurele indicatoren toegelaten voor zover de indirecte effecten via opvoedingswaarden geen verklaring konden bieden. In totaal hebben we 50 van de oorspronkelijk 126 gefixeerde parameters vrijgelaten, waarvan er 20 significant op het .05 niveau zijn en geen enkele op het .01 niveau (zie Tabel 3).

In het nu volgende zullen de resultaten van de LISREL-analyses besproken worden. Er zij opgemerkt dat in Figuur 2 uitsluitend sterk significante effecten ($\alpha \leq .01$) zijn opgenomen. Zie voor een volledig overzicht van de geconstateerde samenhangen en effecten: Tabel 2 tot en met 4 in de Appendix.



$\chi^2 = 150.60$; $df = 124$ ($p = .05$);
 Goodness of fit (GFI) = .99; Root Mean Square Residual (RMR) = .03;
 $R^2 = .26$

Figuur 2. Empirisch model voor de verklaring van opvoedingswaarden en opvoedingsgedragingen zowel voor moeders ($N = 570$) als voor vaders ($N = 647$). Gestandaardiseerde effecten

Causale analyses. Uitkomsten

In overeenstemming met de Kohn-these blijken opvoedingswaarden een nagenoeg volledige interpretatie te kunnen geven voor het effect van sociaal-structurele factoren op ouderlijk opvoedingsgedrag. Zie Figuur 2⁴. Tabel 2 geeft een volledig overzicht van de γ -waarden.

Zoals gezegd hebben we met de 'multi-sample' optie van LISREL onderzocht in hoeverre de γ - en β - en ψ -waarden binnen het vader-model significant afwijken van die binnen het moeder-model. Daarbij bleek dat voor wat betreft deze waarden het vader-model

bijna volstrekt identiek was aan het moeder-model. Het aan elkaar gelijkstellen van de betreffende parameters leverde geen significante verslechtering van de 'fit' van het model op (verschil in $\chi^2 = 50.07$ bij verschil van $df = 48$)⁵. Dit ondanks het feit dat de intercorrelaties van de sociaal-structurele variabelen en de arbeidscondities bij de vaders duidelijk verschillen van die bij de moeders⁶.

In tegenspraak met Kohns theorie en bevindingen is dat niet sociale klasse en de daarmee samenhangende arbeidscondities verantwoordelijk kunnen worden gesteld voor de mate waarin ouders in hun opvoeding gericht zijn op

conformiteit maar dat het opleidingsniveau de belangrijkste invloed uitoefent. Naarmate ouders een hogere opleiding hebben genoten zijn zij minder gericht op de opvoedingswaarden aanpassing en gehoorzaamheid in de opvoeding van hun kinderen. Dat opleiding en dus niet sociale klasse, arbeidscondities of zelfs het al dan niet betaald werkzaam zijn ten grondslag ligt aan de ontwikkeling van de ouderlijke opvoedingswaarden conformiteit en gehoorzaamheid is volgens ons de belangrijkste reden dat we voor vaders en moeders vrijwel identieke verklaringsmodellen hebben gevonden.

In overeenstemming met Kohns theorie is dat de mate waarin ouders gericht zijn op conformiteit – althans gemeten met behulp van de door hem ontwikkelde conformiteitsindex – een sterk negatief effect heeft op de mate waarin zij bij hun kinderen onafhankelijkheid en zelfverantwoordelijkheid benadrukken en aanmoedigen, terwijl tegelijkertijd geen sterk significante effecten zijn te bespeuren op de mate waarin ouders affectie tonen en hun kinderen bestraffen (zie Tabel 3).

In tegenspraak met Kohns theorie gaat er wel degelijk een invloed uit van ouderlijke opvoedingswaarden op opvoedingsgedragingen als affectie en straf. Gemeten met alternatieve beoordelingsschalen blijkt de opvoedingswaarde conformiteit in relatief sterke mate gedrag als straf te beïnvloeden. En de opvoedingswaarde persoonlijke autonomie beïnvloedt in relatief sterke mate de ouderlijke gedragingen autonomie, affectie en straf.

Opmerkelijk daarbij is dat deze alternatieve operationalisaties van conformiteit en zelfbepaling geen sterke onderlinge samenhang tonen, terwijl Kohns conformiteitsindex alleen positief samenhangt met de alternatieve conformiteitsschaal (zie Tabel 4). Dit houdt in dat deze maten verschillende aspecten van conformiteit en autonomie tot uitdrukking lijken te brengen⁷. Kohns conformiteitsindex drukt vooral conformiteit aan maatschappelijke conventies uit, terwijl de alternatieve maat voor conformiteit meer nadrukkelijk gehoorzaamheid aan de regels van het gezin indiceert. Daarnaast springt in het oog dat de alternatieve maat voor conformiteit wel maar de alternatieve maat voor autonomie niet beïnvloed wordt door sociaal-structurele factoren⁸. De

Witte (1990) vond vergelijkbare resultaten: de conformiteitsitems van Kohn bleken wel en de zelfbepalings-items van Kohn bleken niet samen te hangen met sociale klasse.

In antwoord op de eerste vraag van dit onderzoek kan vastgesteld worden dat Kohns sociaal-structurele handelingstheorie – zelfs in sterk gereviseerde vorm – geen voldoende verklaring kan geven voor het tot stand komen van ouderlijke opvoedingswaarden. Het zijn niet de sociale klasse en de arbeidscondities maar het opleidingsniveau dat de opvoedingswaarden beïnvloedt en dan met name de conformiteitswaarden daarbinnen. De opvoedingswaarde persoonlijke autonomie, welke in tegenstelling tot conformiteit, van invloed is op *alledrie* de hier gemeten opvoedingsgedragingen, kan zelfs in het geheel niet worden verklaard vanuit sociaal-structurele kenmerken.

Teruggrijpend op de overige, in paragraaf 3 gestelde vragen kan het volgende geconcludeerd worden.

- * Opleiding is van groter belang voor de verklaring van ouderlijke opvoedingswaarden dan sociale klasse.
- * Arbeidsomstandigheden hebben geen direct effect op ouderlijke opvoedingswaarden.
- * Het empirisch verklaringsmodel voor ouderlijk opvoedingshandelen van de vaders wijkt nauwelijks af van dat van de moeders. Voor zover we verschillen vinden zijn die niet toe te schrijven aan uiteenlopende arbeidsomstandigheden maar aan uiteenlopende effecten van opleiding voor het tonen van affectie naar de kinderen.
- * Kohns conformiteitsindex toont overkomstig de verwachtingen een sterk negatieve samenhang met de mate waarin ouders onafhankelijk gedrag bij hun kinderen waarderen. Deze negatieve samenhang wordt echter gerelativeerd door de lage betrouwbaarheid van Kohns conformiteitsindex.
- * De hier gebruikte maten voor ouderlijk opvoedingshandelen zijn zeer betrouwbaar, verwijzen naar daadwerkelijk gedrag en zijn valide gelet op de verwachte samenhangen met opvoedingswaarden.

6 Discussie

Op grond van deze resultaten kunnen enkele algemene conclusies getrokken worden.

Ten eerste, van de vijf in het onderzoek betrokken sociaal-structurele factoren bleek alleen opleiding sterk significante effecten te hebben op de ouderlijke conformiteitswaarden. De opvoedingswaarde met de meest brede invloed op ouderlijk gedrag, te weten: persoonlijke autonomie kan evenwel niet door opleiding verklaard worden. De twee factoren inzake arbeidscondities, te weten: arbeidssatisfactie en werkstress, blijken geen sterk significante effecten op opvoedingswaarden en opvoedingsgedragingen te vertonen.

Een praktische consequentie hiervan – maar dit terzijde – is dat bij onderzoek naar ouderlijke opvoedingsaspecten veel gemakkelijker de gegevens van de moeders in het onderzoek betrokken kunnen worden: als namelijk de genoten opleiding het meest bijdraagt in de verklaring, hoeft niet gezocht te worden naar vervangende gegevens voor partners die geen betaald werk buitenshuis verrichten (Ganzeboom & De Graaf, 1989).

Veel belangrijker dan deze praktische consequentie is dat Kohns sociaal-structurele handelingstheorie op twee fundamentele punten aan revisie toe is. Immers niet het actuele beroep en de daarmee samenhangende arbeidscondities bepalen op welke wijze mensen de werkelijkheid zien en interpreteren en verschillende visies ontwikkelen omtrent wat waardevol en wenselijk is maar voornamelijk de opleiding die zij vaak ver voor de uitoefening van hun huidige werkzaamheden (betaald dan wel onbetaald) hebben genoten.

De vragen die naar aanleiding van deze uitkomsten verder onderzoek stimuleren zijn:

- a) *Wat maakt het dat de factor opleiding zo'n centrale betekenis heeft?*
- b) *Welke in dit onderzoek ontbrekende factor kan verantwoordelijk worden gesteld voor de opvoedingswaarde persoonlijke autonomie?*

ad a. In de literatuur zijn uiteenlopende verklaringen te vinden voor de invloed van de factor opleiding. Zoals we reeds in het begin opmerkten heeft Kohn zelf al een aanzet gegeven: in lagere opleidingen alsook tijdens de eerste

schooljaren zouden leerlingen vooral geconfronteerd worden met de waarde conformiteit terwijl ze in latere schooljaren en in hogere opleidingen vooral tot zelfbeschikking en autonomie zouden worden aangezet (Kohn, 1983). Bowles en Gintis (1976) en Laosa (1982) kwamen al tot vergelijkbare interpretaties door te veronderstellen dat langduriger geschoolden in sterkere mate dan geschoolden met een korte (en veelal lagere) opleiding gesocialiseerd zijn in op onderzoek en ondersteuning gebaseerde onderwijsmethoden. En zij zullen die aangeleerde methoden ook meer nadrukkelijker hanteren in de opvoeding van hun kinderen. Ouders met een kortere opleiding zouden vooral het dwingende voorbeeldkarakter in hun opvoeding benadrukken.

Terwijl de voorafgaande auteurs dus vooral het socialiserend effect als toegevoegde waarde aan de opleiding benadrukken voor de totstandkoming van ouderlijke opvoedingswaarden zijn andere auteurs op zoek gegaan naar effecten die na de opleiding optreden. Effecten die wellicht een interpretatie kunnen aandragen voor de massieve invloed van opleiding. Zo komen Gerris en Janssens (1987) en Janssens en Gerris (1988), zich baserend op Bernsteins socialisatietheorie, tot de bevinding dat de culturele oriëntatie binnen het gezin een gedeeltelijke interpretatie kan geven voor het verband tussen sociaal milieu en opvoedingsoriëntaties. Hoewel de auteurs geen rechtstreeks effect van opleiding op de culturele oriëntatie rapporteren, is de constatering niet zonder belang. Uit het, sterk door Bourdieu's stratificatietheorie geïnspireerde, onderzoek van De Graaf (1987) kwam namelijk naar voren dat met behulp van de culturele oriëntatie van de ouders de effecten van hun opleiding in belangrijke mate geïnterpreteerd konden worden. Daarnaast zij vermeld dat Gerris et al. (1991a; 1991b), startend vanuit de cognitief-structurele theorie van Newberger, tot de conclusie kwamen dat de mate van 'complexity of reasoning' een 'belangrijke intermediaire variabele was tussen sociaal milieu en ouderlijke opvoedingswaarden.

Ten slotte moet in dit verband gewezen worden op de mogelijkheid dat de invloed van opleiding een schijneffect kan zijn. Niet het opleidingsniveau van de ouders maar het sociale milieu waarin de ouders zelf zijn opge-

groeid zou verantwoordelijk zijn voor de totstandkoming van ouderlijke opvoedingswaarden en -gedragingen (zie voor een genuanceerde bekrachtiging van deze hypothese: McKinley, 1964).

Samenvattend kunnen dus drie elkaar niet uitsluitende hypothesen geformuleerd worden ten aanzien van het belang van de factor opleiding met betrekking tot opvoedingswaarden: opleiding impliceert een socialisatie-effect, opleiding heeft een indirect effect en opleiding heeft een schijneffect. Nader onderzoek moet leren welke verklaring het meest plausibel is.

ad b. Zoals we reeds stelden is er nog een tweede reden waarom Kohns theorie aan revisie toe is. Uit onze resultaten bleek namelijk dat de ouderlijke opvoedingswaarde persoonlijke autonomie noch verklaard kon worden met behulp van sociaal-structurele factoren noch met factoren die verwijzen naar arbeidscondities.

Kohn (1969/1989) kwam tot de conclusie dat de 'religieuze factor' in vergelijking met de sociaal-structurele factor irrelevant is voor de verklaring van ouderlijke opvoedingswaarden en -gedragingen. Recent onderzoek toont aan dat Kohns conformiteitsindex een tamelijk sterke samenhang toont met een christelijke geloofsovertuiging (Felling, Peters & Schreuder, 1991). Dit wijst er op dat zelfs een sterk gereviseerde sociaal-structurele handelingstheorie te beperkt van opzet is om ouderlijke opvoedingswaarden en -gedragingen op een adequate wijze te verklaren. Omdat ook andere auteurs religieuze factoren op het spoor zijn gekomen voor de verklaring van de opvoeding (bijv. Alwin, 1986; Dijkstra & Van Laarhoven, 1990; Ellison & Sheekat, 1993), willen we in nader onderzoek een cultureel model voor ouderlijk handelen ontwikkelen en toetsen in de stellige verwachting daarmee enkele antecedenten van de ouderlijke opvoedingswaarde persoonlijke autonomie te achterhalen.

Ten tweede: in dit onderzoek is uitgebreid aandacht besteed aan de betrouwbaarheid en validiteit van de door Kohn ontwikkelde conformiteitsindex. Daarbij bleek de betrouwbaarheid van de index als twijfelachtig te moeten worden beoordeeld, wat inhoudt dat de vraag naar de geldigheid van Kohns conformiteitsindex eigenlijk zinloos is: zonder voldoende be-

trouwbaarheid kan er van validiteit immers geen sprake zijn. Voor wat betreft de betrouwbaarheid het volgende. Een kwart van de respondenten kon niet adequaat de aangeboden opvoedingswaarden rangschikken wat op zijn minst vragen opwerpt omtrent de foutenmarge in de gegevens van degenen die daartoe wel in staat bleken. Verder zijn rangorde-data niet zonder meer te analyseren met behulp van gangbare technieken zoals factoranalyse en Likert-analyse. Uit aangepaste analyses bleek evenwel dat de betrouwbaarheid van de Kohn-index laag is.

LISREL-analyses lieten een relatief sterke samenhang zien tussen Kohns conformiteitsindex en de alternatieve en veel betrouwbaardere conformiteitsschaal. Ondanks deze relatief sterke samenhang bleken deze schalen geen vergelijkbare samenhangen met ouderlijk opvoedingsgedragingen te vertonen. Kohns conformiteitsindex bleek vooral negatief samen te hangen met het ouderlijke gedrag autonomie, wat een ondersteuning inhoudt van Kohns opvattingen. Aan deze sterke samenhang moet echter niet al te veel gewicht worden toegekend gezien het onbetrouwbare karakter van de Kohn-index. De alternatieve conformiteitsmaat liet vrijwel uitsluitend haar invloed gelden op straffen in overtredingssituaties. De ouderlijke opvoedingswaarde persoonlijke autonomie bleek noch een negatieve samenhang te vertonen met de alternatieve conformiteitsmaat noch met Kohns conformiteitsindex, wat in duidelijke tegenspraak is met Kohns veronderstellingen. Met behulp van de ouderlijke opvoedingswaarde persoonlijke autonomie kon daarentegen een veel breder spectrum aan gedragsuitingen verklaard worden. Dit laatste is niet zonder belang omdat met gebruikmaking van uitsluitend de Kohn-index gemakkelijk het onjuiste beeld zou ontstaan dat de opvoedingswaarden die aan ouderlijk gedrag ten grondslag liggen geen invloed zouden hebben op basale gedragingen als affectie en straf. Op grond van deze gegevens lijkt het daarom aan te bevelen om de Kohn-items niet als te rangschikken items maar als Likert-items voor te leggen (cf. Jansen, 1983; De Witte, 1990).

Als afsluiting willen we enkele suggesties voor nader onderzoek kort aanduiden. Op de eerste plaats impliceert het feit dat we geen directe invloed van arbeidscondities op ouder-

lijke opvoedingswaarden nog niet dat de werk-omstandigheden volstrekt irrelevant zouden zijn. Wellicht behoeft het hier gepresenteerde empirisch model nadere specificatie. Uit onderzoek zou kunnen blijken dat bijvoorbeeld het al dan niet zelfstandig werkzaam zijn als een moderator-variabele fungeert: bij zelfstandigen zouden de hier gevonden effecten veel sterker zijn dan bij personen werkzaam in loondienst (Miller & Swanson, 1958). Daarnaast zouden ook gezinssituationele aspecten zoals gezinssfeer en huwelijkssatisfactie, of kenmerken zoals temperament, leeftijd en geslacht tot een verdere specificatie van het model kunnen leiden.

Ten tweede zou onderzocht kunnen worden wat de effecten zijn van de ouderlijke opvoeding op karakteristieken van kinderen. Uit onderzoek komt naar voren dat niet zozeer de door ouders gerapporteerde gedragingen en overtuigingen van invloed zijn op de gedragingen en overtuigingen van hun kinderen maar vooral de gedragingen en overtuigingen zoals ze door hun kinderen worden waargenomen (Acock & Bengtson, 1980; Van der Slik, 1992). Hiermee samenhangend zou eveneens onderzocht kunnen worden in hoeverre ouderlijke opvoedingswaarden en -gedragingen van invloed zijn op de opleidingslooptijden van kinderen (De Graaf, 1987; Meijnen, 1977; Morgan, Alwin & Griffin, 1979).

Noten

- 1 Er moet benadrukt worden dat Jansen (1983) een conformiteitsfactor en een daarmee positief correlerende zelfbepalingsfactor heeft achterhaald. Dit resultaat wordt door hem toegeschreven aan het feit dat hij de Kohn-items niet als te rangschikken waarden heeft gepresenteerd maar als Likert-statements heeft voorgelegd. Dat Kohn een index vindt waarin conformiteit tegenover zelfbepaling staat, is volgens Jansen een artefact van de gevolgde procedure.
- 2 Voor zover wij weten heeft tot nu toe geen enkele onderzoeker de betrouwbaarheid van de Kohn-index gerapporteerd zodat een vergelijking niet mogelijk is.
- 3 Nadere inspectie van de antwoordpatronen deed vermoeden dat ouders vaak grote moeite hadden om een rangorde aan te brengen in de

aangeboden opvoedingswaarden, gezien het feit dat een kwart van de ouders geen volledig scorepatroon liet zien. Jansen (1983, p. 40) vond vergelijkbare resultaten. De wens van meer handzame alternatieven lijkt achteraf gezien geen overbodige luxe te zijn geweest.

- 4 We hebben om de stabiliteit van dit empirisch model te testen diverse steekproeven uit de gezinnen getrokken en de effecten opnieuw geschat. De uitkomsten weken niet noemenswaardig af van de hier gepresenteerde resultaten.
- 5 We vonden slechts één markant verschil: vaders met een hogere opleiding tonen *meer* affectie naar hun kinderen dan vaders met een lagere opleiding ($\gamma = .08$; $\alpha \leq .05$); moeders met een hogere opleiding daarentegen tonen juist *minder* genegenheid naar hun kinderen dan moeders met een lagere opleiding ($\gamma = -.10$; $\alpha \leq .05$). Hoewel intrigerend, wachten we met een verklaring voor dit verschil tot deze bevindingen gerepliceerd worden.
- 6 De tabel met intercorrelaties tussen de onafhankelijke variabelen is verkrijgbaar bij de eerstgenoemde auteur.
- 7 We hebben om deze claim te onderzoeken twee extra analyses uitgevoerd. Een waarbij alleen Kohns conformiteitsindex is ingebracht en een analyse waarbij alleen de alternatieve schalen voor conformiteit en persoonlijke autonomie zijn betrokken. De uitkomsten weken in zoverre af van de hier gepresenteerde gegevens dat Kohns conformiteitsindex een zwak positief effect had op het geven van straf terwijl de alternatieve conformiteitsschaal een zwak negatief effect had op de gedragsuiting autonomie. Gezien de samenhang tussen deze conformiteitsschalen is dat niet verrassend.
- 8 Uiteindelijk waren dit voor ons doorslaggevende argumenten om de drie opvoedingswaarden-schalen als drie afzonderlijke factoren in het empirische model in te brengen en niet als verwijzend naar één onderliggend construct.

Literatuur

- Acock, A.C., & Bengtson, V.L. (1982). Socialization and attribution processes: actual versus perceived similarity among parents and youth. *Journal of Marriage and the Family*, 42, 501-515.

- Alwin, D.F. (1986). Religion and parental child-rearing orientations: evidence of a Catholic-Protestant convergence. *American Journal of Sociology*, 92, 412-440.
- Belsky, J. (1981). Early human experience: a family perspective. *Developmental Psychology*, 17, 3-23.
- Belsky, J. (1984). The determinants of parenting: a process model. *Child Development*, 55, 83-96.
- Belsky, J., Robins, E., & Gamble, W. (1984). The determinants of parental competence: toward a contextual theory. In M. Lewis (Ed.), *Beyond the dyad* (pp. 251-280). New York: Plenum Press.
- Blau, P., & Duncan, O.D. (1967). *The American occupational structure*. New York: Wiley & Sons.
- Bosker, R., Meijnen, G.W., & Velden, R. van der (1988). Sociale stratificatie en maatschappelijke oriëntaties. *Mens & Maatschappij*, 63, 156-174.
- Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: educational reform and the contradictions of economic life*. New York: Basic Books.
- Brock, A.J.L.L. de (1994). *Ouderlijk opvoedkundig handelen: de invloed van ouder-, kind- en contextuele kenmerken*. Nijmegen: Universiteitsdrukkerij Nijmegen.
- Deković, M. (1991). *The role of parents in the development of child's peer acceptance*. Nijmegen: Institute of Family Studies, Katholieke Universiteit.
- Dijkstra, A.B., & Laarhoven, P. van (1990). Een religieuze factor in opvoedingsstijlen? *Sociologische Gids*, 5, 320-332.
- Ellison, Ch. G., & Sheekat, D.E. (1993). Obedience and autonomy: religion and parental values reconsidered. *Journal for the Scientific Study of Religion*, 32, 313-329.
- Erikson, R., Goldthorpe, J.H., & Portocarero, L. (1979). Intergenerational class mobility in three West European Countries: England, France and Sweden. *The British Journal of Sociology*, 10, 415-441.
- Felling, A., Peters, J., & Schreuder, O. (1991). *Dutch religion. The religiousness of the Netherlands after the cultural revolution*. Nijmegen: ITS.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior. An introduction to theory and research*. Reading, MA: Addison-Wesley Publishing Company.
- Ganzeboom, H.B.G., & Graaf P.M. de (1989). Inter-generationale opleidingsmobiliteit in Nederland van geboortecohorten 1891-1960. *Sociale Wetenschappen*, 32, 263-278.
- Gecas, V., (1979). The influence of social class on socialization. In W.R. Burr, R. Hill, F.I. Nye & I.L. Reiss (Eds.), *Contemporary theories about the family. Research-based theories. Vol. 1* (pp. 365-404). New York: The Free Press.
- Gerris, J.R.M., Boxtel, D.A.A.M. van, Vermulst, A.A., Janssens, J.M.A.M., Zutphen, R.A.H. van, & Felling, A.J.A. (1992). *Child rearing and family in the Netherlands. Documentation of a national representative survey of child-rearing, family relations and family processes in 1990*. Nijmegen: Institute of Family Studies, University of Nijmegen.
- Gerris, J.R.M., Vermulst, A.A., Boxtel, D.A.A.M. van, Janssens, J.M.A.M., Zutphen, R.A.H. van, & Felling, A.J.A. (1993). *Parenting in Dutch families. A representative description of Dutch family life in terms of validated concepts representing characteristics of parents, children, the family as a system and parental socio-cultural value orientations*. Nijmegen: Institute of Family Studies, University of Nijmegen.
- Gerris, J.R.M., Deković, M., & Janssens, J.M.A.M. (1991a). Basisoriëntaties van ouders op kind en opvoeding; een uitbreiding van het model Kohn. *Gezin*, 3, 216-233.
- Gerris, J.R.M., Deković, M., & Janssens, J.M.A.M. (1991b). Cultuuroverdracht in de opvoeding. Een uitbreiding van het model Kohn. In A.J.A. Felling & J. Peters (Red.), *Cultuur en sociale wetenschappen. Beschouwingen en empirische studies* (pp. 199-226). Nijmegen: ITS.
- Gerris, J.R.M., & Janssens, J.M.A.M. (1987). Opvoedingsgedrag van ouders in overtredingssituaties: een onderzoek naar beïnvloedende factoren. *Pedagogische Studiën*, 64, 295-307.
- Gerris, J.R.M., Vermulst, A.A., & Brock, A.J.L.L. de (1991). Kwaliteit partnerrelatie, overeenstemming tussen ouders, en geslachtsspecifieke opvoedingspatronen in gezinsdyades: een exploratie. In J.R.M. Gerris (Red.), *Ouderschap en ouderlijk functioneren. Gezinsonderzoek, nr4* (pp. 83-108). Amsterdam/Lisse: Swets & Zeitlinger.
- Gerris, J.R.M., Vermulst, A.A., & Siebenheller, F.A. (1990). Overeenstemming tussen ouders: onderliggende dimensies en samenhangen met opvoedingsoriëntaties en opvoedingsgedragingen van vaders en moeders afzonderlijk. In J.R.M. Gerris & J. Van Acker (Red.), *Gezin: aspecten van sociale ondersteuning en intern functioneren* (pp. 147-173). Amsterdam/Lisse: Swets & Zeitlinger.

- Graaf, P. M. de (1987). *De invloed van financiële en culturele hulpbronnen in onderwijsloopbanen*. Nijmegen: ITS.
- Haaf, P. G. J. ten (1993). *Opvoedingsdimensies: convergente en discriminante validiteit*. Nijmegen: UDN.
- Haaf, P. G. J. ten, Janssens, J. M. A. M., & Gerris, J. R. M. (1994). Child-rearing measures: convergent and discriminant validity. *European Journal of Psychological Assessment*, 10, 111-128.
- Haaf, P. G. J. ten, & Janssens, J. M. A. M. (1991). Confirmerende en discriminerende validiteit van ouderlijk opvoedingsgedrag. *Gezin*, 3, 119-135.
- Jansen, H., (1983). Klasse- en sexe-specificiteit in de waardering van conformiteit en zelfbepaling. Enkele ervaringen met de schaal van Kohn. *Sociologische Gids*, 30, 37-44.
- Janssens, J. M. A. M., & Gerris, J. R. M. (1988). Sociaal milieu en reacties van ouders op disciplinerings-situaties: een empirisch verklaringsmodel. *Pedagogische Studiën*, 65, 185-197.
- Jöreskog, K., & Sörbom, D. (1989). *LISREL 7: a guide to the program and applications*. Chicago, Gornichem: SPSS Inc.
- Jöreskog, K., & Sörbom, D. (1990). *SPSS LISREL 7 and PRELIS. User's guide and reference*. Chicago, IL: SPSS Inc.
- Kirkpatrick, L. A., & Shaver, Ph. R. (1990). Attachment theory and religion: childhood attachments, religious beliefs, and conversion. *Journal for the Scientific Study of Religion*, 29, 315-334.
- Kohn, M. L., (1963). Social class and parent-child relationships: an interpretation. *American Journal of Sociology*, 68, 471-480.
- Kohn, M. L., (1976). Social class and parental values: another confirmation of the relationship. *American Sociological Review*, 41, 538-545.
- Kohn, M. L., (1983). On the transmission of values in the family: a preliminary formulation. In A. C. Kerckhoff (Ed.), *Research in Sociology of Education and Socialization. A Research Annual. Personal change over the life course. Vol. 4* (pp. 3-14). Greenwich/London: JAI Press.
- Kohn, M. L., (1989 [1969]). *Class and Conformity. A study in values. With a reassessment, 1977*. Chicago, IL: The University of Chicago Press.
- Kohn, M. L., Naoi, A., Schoenbach, C., Schooler C., & Slomczynski, K. M. (1990). Position in the class structure and psychological functioning in the United States, Japan, and Poland. *American Journal of Sociology*, 95, 964-1008.
- Laosa, L. M., (1982). School, occupation, culture, and family: the impact of parental schooling on the parent-child relationship. *Journal of Educational Psychology*, 74, 791-827.
- Liere, C. van, & Winkels, J. W. (1987). Zelfbepaling en conformiteit in de opvoeding: veranderde waarden tussen 1972 en 1986. *M&T Bulletin* 74. Sociaalwetenschappelijk Instituut Groningen.
- McKinley, D. G., (1964). *Social class and family life*. New York: The Free Press of Glencoe.
- Meijnen, G. W., (1977). *Maatschappelijke achtergronden van intellectuele ontwikkelingen. Een empirisch onderzoek naar de invloed van de statusspecifieke opvoeding op de ontwikkeling van intelligentie en leerprestaties*. Groningen: Wolters-Noordhoff.
- Miller, D. R., & Swanson, G. E. (1958). *The changing American parent: a study in the Detroit area*. New York: Wiley & Sons.
- Morgan, W. R., Alwin, D. F., & Griffin, L. J. (1979). Social origins, parental values, and the transmission of inequality. *American Journal of Sociology*, 85, 156-166.
- Rollins, B. C., & Thomas, D. L. (1979). 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. Research-based theories* (pp. 317-364). London: Free Press.
- Siebenheller, F. A., (1990). *Problematische opvoedingssituaties: percepties, emoties en disciplineringsreacties van ouders*. Nijmegen: Institute of Family Studies, Katholieke Universiteit.
- Slik, F. W. P. van der, (1992). *Overtuigingen, attitudes, gedrag en ervaringen. Een onderzoek naar de godsdienstigheid van ouders en van hun kinderen*. Dissertatie, Nijmegen.
- Slomczynski, K. M., Miller, J., & Kohn, M. L. (1981). Stratification, work, and values: a Polish-United States comparison. *American Sociological Review*, 46, 720-744.
- Vermulst, A. A., Gerris, J. R. M., Franken, W. M., & Janssens, J. M. A. M. (1986). Determinanten van ouderlijk functioneren tegen de achtergrond van de theorie van Kohn. In J. R. M. Gerris (Red.), *Pedagogisch onderzoek in ontwikkeling. Opvoedingsprocessen, gezinsproblemen, leren, diagnostiek en probleemgedrag* (pp. 19-42). Nijmegen: ITS.

Vermulst, A. A., Gerris, J. R. M., & Siebenheller, F. A. (1987). Opvoedingsdoelen: bepaling van de onderliggende betekenisdimensies in het meetinstrument opvoedingsdoelen met behulp van een drietal multivariate analysetechnieken. In J. R. M. Gerris & J. Van Acker (Red.), *Gezin: Onderzoek en Hulpverlening* (pp. 67-85). Lisse: Swets & Zeitlinger.

Williamson, R. L., (1984). A partial replication of the Kohn-Gecas-Nye thesis in a German sample. *Journal of Marriage and the Family*, 46, 971-979.

Witte, H. de (1990). *Conformisme, radicalisme en machteloosheid. Een onderzoek naar de sociaal-culturele en sociaal-economische opvattingen van arbeiders in Vlaanderen*. Leuven: Hoger Instituut voor de Arbeid.

Wright J. D., & Wright, S. R. (1976). Social class and parental values for children: A partial replication and extension of the Kohn thesis. *American Sociological Review*, 14, 527-537.

Manuscript aanvaard 24-10-1995

Auteurs

F. W. P. van der Slik is postdoctoraal onderzoeker bij de Faculteit Maatschappijwetenschappen van de Katholieke Universiteit Nijmegen.

J. R. M. Gerris is hoogleraar algemene pedagogiek aan de Katholieke Universiteit Nijmegen.

A. J. A. Felling is hoogleraar methodenleer aan dezelfde universiteit.

Correspondentie-adres: F. W. P. van der Slik, Katholieke Universiteit Nijmegen, Postbus 9104, 6500 HE Nijmegen. E-mail: F.vanderSlik@MAW.KUN.NL

Abstract

Educational background of parents and their child-rearing. A study within the Kohn-tradition

F. W. P. van der Slik, J. R. M. Gerris & A. J. A. Felling.
Pedagogische Studiën, 1996, 73, 338-356.

Examining the relationships between social-structural conditions of parents, their child-rearing values and their child-rearing practices, an extensive theoretical model was tested. Data of a representative sample of Dutch families ($N=788$) were used. Within these families the target-child was between 9 and 16 years old. Causal analyses, by means of structural equation models with a multigroup approach, indicated that the empirical models for mothers and fathers are similar. A second major outcome was that only parental education appears to have substantive effects on two child-rearing values: 'conformity to external authority' (Kohn) and 'conformity to parental rules' (Gerris), while parental education appears to have no effect on the third child-rearing value: 'personal autonomy'. Effects of other social-structural conditions are virtually absent. Additionally, it showed that conformity to external authority has a negative effect on stimulating the child to behave autonomously, while conformity to parental rules has a strong effect on parent's use of punishment. Finally, parents who value personal autonomy, express affection to the child, stimulate its autonomy, and punish it less often. Several explanations of the massive effect of parental educational background are discussed.

Appendix

Tabel 2

Effecten van sociaal-structurele indicatoren en arbeidscondities op opvoedingswaarden en -gedragingen (γ -matrix)^a

VADERS/ MOEDERS	Klasse 1	Klasse 3	Klasse 4	Klasse 5	Klasse 6	Opleiding	Gezins- inko- men	Ar- beids- duur	Wer- kend	Ar- beids- satis- factie	Werk- stress
Kohns conformiteitsindex	0	0	.02 ^c	-.01 ^c	.12* ^c	-.38**	0	-.00 ^c	0	.12* ^c	0
Persoonlijke autonomie	0	0	.00 ^c	.04 ^c	-.02 ^c	-.03	0	-.02 ^c	0	.04 ^c	0
Conformiteit	0	0	.05 ^c	.05 ^c	.09* ^c	-.30**	0	-.03 ^c	-.05 ^c	.09* ^c	0
Autonomie	0	0	0	-.09* ^c	0	.08* ^c	0	.09* ^c	0	.05 ^c	0
Affectie-expressie	0	0	-.06 ^c	0	-.10* ^c	.08/- .10 ^{b,c}	0	0	0	-.10* ^c	0
Straf	0	0	0	.06 ^c	0	0	0	0	0	0	0

*: $\alpha \leq .05$; **: $\alpha \leq .01$; anders: n.s.; 0: gefixeerd

Klasse 2 (lagere leidinggevenden) is referentiecategorie

a: als er maar één numerieke waarde in de tabel staat betekent dit dat er geen verschil is tussen vaders en moeders

b: linkerwaarde voor de vaders ($\alpha \leq .05$); rechterwaarde voor de moeders ($\alpha \leq .05$)

c: oorspronkelijk gefixeerde parameters

Tabel 3

Effecten van opvoedingswaarden op opvoedingsgedragingen (β -matrix)^a

VADERS/MOEDERS ^a	Opvoedingswaarden			Opvoedingsgedragingen		
	Kohns conformiteitsindex	Persoonlijke Autonomie	Conformiteit	Autonomie	Affectie-expressie	Straf
Kohn's conformiteitsindex	0	0	0	0	0	0
Persoonlijke autonomie	0	0	0	0	0	0
Conformiteit	0	0	0	0	0	0
Autonomie	-.33**	.31**	.10*	0	0	0
Affectie-expressie	.10*	.32**	-.01	0	0	0
Straf	.01	-.22**	.48**	0	0	0

*: $\alpha \leq .05$; **: $\alpha \leq .01$; anders: n.s.; 0: gefixeerd

a: als er maar één numerieke waarde in de tabel staat betekent dit dat er geen verschil is tussen vaders en moeders

Tabel 4

Correlaties tussen residuele fouten van de afhankelijke variabelen (Ψ -matrix)^a

VADERS/MOEDERS ^a	Opvoedingswaarden			Opvoedingsgedragingen		
	Kohns conformiteitsindex	Persoonlijke Autonomie	Conformiteit	Autonomie	Affectie-expressie	Straf
Kohn's conformiteitsindex	.81**					
Persoonlijke autonomie	-.13*	.99**				
Conformiteit	.40**	.12*	.86**			
Autonomie	0	0	0	.75**		
Affectie-expressie	0	0	0	.27**	.88**	
Straf	0	0	0	.14**	.18**	.74**

*: $\alpha \leq .05$; **: $\alpha \leq .01$; anders: n.s.; 0: gefixeerd

a: als er maar één numerieke waarde in de tabel staat betekent dit dat er geen verschil is tussen vaders en moeders

Retentie-effecten van een nascholingsprogramma voor effectieve instructie en klasmanagement*

S. Veenman en J. Raemaekers**

Samenvatting

In deze studie worden de lange-termijn effecten van een nascholingsprogramma voor effectieve instructie en klasmanagement beschreven. In twee eerdere studies werd ingegaan op de korte-termijn effecten. De eerste evaluatiestudie werd uitgevoerd in het schooljaar 1986/87, de tweede in 1989/90. In de tweede evaluatiestudie werd ook het effect van coaching beproefd. Eind 1992 werden de lange-termijn of retentie-effecten van het nascholingsprogramma onderzocht. Dit geschiedde door middel van een 'untreated control group design' met een voormeting, een nameting en een retentiemeting. De gegevens betreffende de vergelijking tussen de experimentele groep en de controlegroep lieten zien dat het nascholingsprogramma een significant effect had op de taakgerichtheid van de leerlingen en op aspecten van het instructie- en klasmanagementgedrag van leerkrachten. Een duurzaam effect van coaching werd niet gevonden. Ook werden geen significante verschillen gevonden tussen de experimentele groep die vijf jaar geleden het programma had gevolgd en de experimentele groep die twee jaar geleden het programma had gevolgd. Tussen de nameting en de retentiemeting werden geen significante verschillen aangetroffen. De vooruitgang die direct geboekt werd na de training werd voor de meeste instructie- en klasmanagementvaardigheden vastgehouden, een verdere groei of toename trad niet op. Een effect van het nascholingspro-

gramma op de leerprestaties van de leerlingen werd niet gevonden.

1 Achtergrond

In het basisonderwijs neemt de combinatieklas een prominente plaats in. Volgens de Commissie Evaluatie Basisonderwijs (1994) geeft 53% van de basisschoolleerkrachten les in een combinatieklas. In een combinatieklas geeft de leerkracht les aan twee of meer groepen tegelijkertijd. Op grond van observaties en interviews met leerkrachten uit combinatieklassen werden vijf probleemgebieden geïdentificeerd bij het geven van onderwijs in combinatieklassen, namelijk 1) het efficiënt gebruik van de leertijd, 2) het geven van effectieve instructie, 3) doelmatig klasmanagement, 4) het organiseren van zelfstandig leren, en 5) het schoolbeleid met betrekking tot combinatieklassen (Veenman, Lem, Voeten, Winkelmolen & Lassche, 1986). Op basis van deze vijf probleemgebieden werd een teamgerichte nascholingscursus ontwikkeld, getiteld *Omgaan met combinatieklassen* (Veenman, Lem & Nijssen, 1988)¹. De inhoud van de nascholing was gebaseerd op recente empirische inzichten over goed en succesvol onderwijzen. In twee evaluatiestudies werden de korte-termijn effecten van het nascholingsprogramma vastgesteld. De opzet en de resultaten van deze studies zijn in twee eerdere artikelen in dit tijdschrift beschreven (Veenman, Lem & Roelofs, 1990; Roelofs, Veenman & Raemaekers, 1993) en worden beknopt in paragraaf 4 samengevat. Dit artikel richt zich op de retentie- of lange-termijn effecten van het nascholingsprogramma.

* Dit onderzoek werd uitgevoerd met behulp van een subsidie van het Instituut voor Onderzoek van het Onderwijs (SVO-project 1111).

** Dank zijn wij verschuldigd aan de volgende personen voor hun bijdragen: M. Voeten, C. van den Berg, M. van Moorsel, C. Motshagen, D. Jacobs, N. Hofs en de deelnemende schoolleiders en leerkrachten.

2 Training en transfer

In studies waarin getracht wordt de uitkomsten van trainingen vast te stellen, wordt opgemerkt dat het onderzoek zich tot nu toe weinig gelegen heeft laten liggen aan de transfer-effecten van trainingen (Baldwin & Ford, 1988; Broad & Newstrom, 1992; Simons, 1992). Veel nascholingscursussen leiden niet tot een toepassing van de geleerde inzichten en vaardigheden in de klas als werksituatie (Van Tulder, 1992). De overdracht van het geleerde in een cursus of training naar de werksituatie wordt in de literatuur aangeduid met het begrip 'transfer'. Wanneer deze overdracht is geslaagd is sprake van positieve transfer. Transfer treedt op als de lerende in staat is het geleerde van de trainingssituatie te generaliseren naar de werksituatie en het aldaar gedurende langere tijd toepast. Transfer houdt ook in dat met toenemende oefening op de werkplek het niveau van de toegepaste kennis en vaardigheden na verloop van tijd hoger ligt dan aan het einde van de training.

Gezien de geringe transfer-effecten voor allerlei soorten van trainingen, nemen Broad en Newstrom (1992) aan dat ongeveer 50% van alle trainingsinhouden één jaar na afloop van de training nog toegepast worden. Om de condities voor het optreden van transfer-effecten te versterken stellen zij de vorming van 'transfer partnerships' voor. Dit zijn samenwerkingsverbanden tussen de drie participanten die belang hebben bij de training: de lerenden (de trainees), de trainers (de ontwerpers en de instructeurs van de training) en de managers (de leidinggevendenden in de organisatie die verantwoordelijk zijn voor de toepassing van de trainingsinhouden op de werkplek). Iedere partner kan vóór, tijdens en na de training een aantal maatregelen nemen om het proces van transfer te bevorderen. Door de tijdsdimensie (vóór, tijdens, na de training) te combineren met de rol-dimensie van de betrokken partners (manager, trainer, lerende) hebben Broad en Newstrom (1992) een matrix ontworpen met transferbevorderende maatregelen. Een aantal van de door hen aangereikte suggesties is verwerkt in de opzet en de uitvoering van de nascholing *Omgaan met Combinatieklassen*. Zo werden tijdens een intake-gesprek de leerkrachten op de hoogte gebracht van de doelen, de inhouden, de werkwijzen van de nascholing. Ook werd

gewezen op het nut of de bruikbaarheids-waarde van de inhouden van de nascholing. Zowel de docenten die de nascholing verzorgden als de leerkrachten die de nascholing volgden werden betrokken bij de opzet en uitwerking van het programma. De leerkrachten die aan de nascholing deelnamen hadden allemaal de intentie om hun onderwijs in combinatieklassen te verbeteren. Volgens Den Ouden (1992) vormen de intenties van de lerenden een cruciaal concept bij het aanbrengen en het verklaren van gedragsverandering na een training of opleiding. Gegevens over het feitelijke gedrag van leerkrachten - verzameld via observaties in de klas - werden tijdens de nascholing gebruikt om verbeteringsvoorstellen te formuleren in het licht van de aangeboden informatie. De nascholing was systematisch opgezet en uitgevoerd volgens de door Joyce en Showers (1995) beschreven componenten van effectieve training. In teamverband werden gezamenlijke voorstellen voor verbetering geformuleerd. Oefensituaties om de nieuwe vaardigheden toe te passen werden gecreëerd; tijdens de training via gesimuleerde condities en tijdens het lesgeven via praktijkgeoriënteerde opdrachten. Leerkrachten kregen de mogelijkheid om hun ervaringen met de toepassing van de nieuwe vaardigheden te bespreken met collega's. In gezamenlijk overleg werden lesplannen voorbereid en geëvalueerd. Om transfer te bevorderen werd bovendien voorzien in coaching.

3 Coaching

Studies naar de relatie tussen vormgeving en effecten van nascholing laten zien dat nascholing leidt tot transfer als naast de training ook sprake is van begeleiding of coaching in de klas (Bennett, 1987; Joyce & Showers, 1995). Bij programma's zonder coaching zijn de transfer-effecten minimaal. Coaching wordt door Joyce en Showers (1980) beschreven als: "Hands-on, in-classroom assistance with the transfer and application of skills to the classroom". Aan coaching kunnen vier functies worden onderscheiden: (1) het geven van steun en vertrouwen aan de leerkracht die poogt nieuwe kennis en vaardigheden in zijn klas toe te passen, (2) het geven van inhoudelijke terugkoppeling

over de toepassing van nieuwe kennis en vaardigheden in de klas, (3) het analyseren van de nieuwe kennis en vaardigheden op hun gebruikswaarde in de klas, en (4) het afstemmen van de nieuwe kennis en vaardigheden op de leerlingen.

4 Korte-termijn effecten van het nascholingsprogramma

Zoals gezegd werden in twee evaluatiestudies de korte-termijn effecten van de nascholing vastgesteld. In de eerste evaluatiestudie waren medewerkers van de Vakgroep Onderwijskunde betrokken bij de opzet en uitvoering van de training. In het schooljaar 1986/87 namen 41 leerkrachten van 8 schoolteams aan de nascholing deel (combinaties 2/3 tot en met 7/8). Uit deze groep werden vervolgens 17 leerkrachten geselecteerd voor observaties in de klas (experimentele groep). Negen leerkrachten uit 6 vergelijkbare scholen vormden de controlegroep (zie Tabel 1, Appendix). Zij namen niet aan de nascholing deel. De leerkrachten uit de beide groepen werden direct vóór en na de nascholing tijdens twee rekenlessen en twee lees/taallessen geobserveerd.

In de tweede evaluatiestudie werd op vijf plaatsen in Nederland de nascholing verzorgd door PABO-docenten en schoolbegeleiders. Deze ontvingen een voorscholing om te garanderen dat het programma op eenzelfde wijze werd uitgevoerd als in de eerste studie. In het schooljaar 1989/90 namen 89 leerkrachten van 12 scholen deel aan de nascholing. Uit deze groep werden 28 leerkrachten (groep 3 tot en met 8) geselecteerd voor lesobservaties (experimentele groep). Van deze groep ontvingen 18 leerkrachten naast de nascholing ook nog coaching door de schoolbegeleiders. De controlegroep werd gevormd door 14 leerkrachten van 6 vergelijkbare scholen (zie Tabel 1).

De nascholing omvatte vijf tot zeven bijeenkomsten van elk drie uur, die om de twee weken plaatsvonden. De weken tussen de bijeenkomsten waren bestemd voor het uitvoeren van praktijkopdrachten. De gegevens van de observaties werden naar de leerkrachten teruggelinkt zodanig op grond van deze gegevens en de gegeven theorie verbeteringsplannen konden formuleren. Deze gegevens hadden

vooral betrekking op het taakgerichte gedrag van de leerlingen en op de instructie- en klasmanagementvaardigheden van de leerkrachten.

Uit de resultaten van de eerste evaluatiestudie bleek dat de nascholing effectief was. Er werd een significant effect vastgesteld ten aanzien van het taakgerichte gedrag van de leerlingen en ten aanzien van een aantal instructie- en klasmanagementvaardigheden (Veenman, Lem & Roelofs, 1990). Ook in de tweede evaluatiestudie werd een significant effect van de nascholing vastgesteld op het taakgerichte gedrag van de leerlingen en op het instructie- en klasmanagementgedrag van de leerkrachten. Twee effecten voor coaching werden gevonden, namelijk voor de organisatie van de instructie en voor het omgaan met storingen in de klas. In klassen van gecoachte leerkrachten werd een sterkere vooruitgang in taakgerichte leertijd geconstateerd dan in klassen van niet-gecoachte leerkrachten (Roelofs, Veenman & Raemaekers, 1993).

5 Onderzoeksvragen

Deze studie richt zich op de lange-termijn of retentie-effecten van het nascholingsprogramma *Omgaan met Combinatieklassen*. Het onderzoek kent de volgende vraagstellingen: 1) Passen de leerkrachten die de nascholing in de schooljaren 1986/87 en 1989/90 hebben gevolgd de gewenste instructie- en klasmanagementvaardigheden nog steeds toe? 2) Heeft de nascholing een blijvend effect op het taakgerichte gedrag van de leerlingen? 3) Zijn de effecten van de nascholing groter voor die leerkrachten die ook coaching hebben ontvangen? 4) Heeft de nascholing effect op de leerprestaties van de leerlingen?

6 Opzet

De toetsing van de effecten van nascholing en coaching op het instructie- en klasmanagementgedrag van de leerkrachten geschiedde door middel van een 'untreated control group design' met een voormeting, een nameting en een retentiemeting. Het onderzoek kende een controleconditie ($n = 11$) en twee experimen-

tele condities: niet-gecoachte leerkrachten ($n = 10$) en gecoachte leerkrachten ($n = 8$). Ter vaststelling van het effect van de nascholing en coaching op de leerprestaties van de leerlingen werd gebruik gemaakt van een 'posttest-only design with nonequivalent groups', waarbij de experimentele groep bestond uit 26 klassen en de controlegroep uit 14 klassen.

6.1 Onderzoeksgroep

Oorspronkelijk was het de bedoeling om alleen de scholen uit de tweede evaluatiestudie te benaderen voor deelname aan de retentiestudie. In deze studie werden vóór en na de nascholing niet alleen observaties uitgevoerd, maar werden ook schooltoetsen afgenomen om het effect van de nascholing op de leerprestaties te kunnen vaststellen. In de eerste evaluatiestudie waren geen schooltoetsen afgenomen. De animo van scholen en leerkrachten om aan de retentiestudie deel te nemen was gering. Door de late start van het onderzoek hadden enkele scholen hun prioriteiten anders gelegd. Sommige leerkrachten waren vertrokken. Een aantal leerkrachten wilde vanwege de taakbelasting slechts deelnemen aan een beperkt aspect van de retentiestudie (bijvoorbeeld wel aan de observatiestudie, maar niet aan de afname van schooltoetsen). Een aantal scholen merkte op dat zij vroeger aan de studie hadden deelgenomen omdat zij daar baat bij hadden in de vorm van nascholing; de retentiestudie gaf hun geen profijt. Door deze experimentele uitval, inherent aan elke longitudinale studie, werd besloten ook de leerkrachten te benaderen die de nascholing in het schooljaar 1986/87 hadden gevolgd. Deze uitbreiding gaf de mogelijkheid ook het tijdseffect te onderzoeken: maakten leerkrachten die vanaf 1986/87 met de inhoud van de nascholing werkten, in vergelijking met de leerkrachten die vanaf 1989/90 hiermee werkten, meer gebruik van de gewenste instructie- en klasmanagementvaardigheden?

De samenstelling van de onderzoeksgroepen voor de eerste en tweede evaluatiestudie is beschreven in Veenman, Lem en Roelofs (1990) en in Roelofs, Veenman en Raemaekers (1993). Het navolgende handelt over de onderzoeksgroep van de retentiestudie.

Van de 17 getrainde leerkrachten die deelnamen aan het observatie-onderzoek van de eerste evaluatiestudie uit 1986/87 bleken 5 jaar

later 8 leerkrachten bereid opnieuw deel te nemen aan de observaties voor de retentiestudie. Van de 10 getrainde maar niet-gecoachte leerkrachten die deelnamen aan het observatie-onderzoek van de tweede evaluatiestudie uit 1989/90 bleken 2 jaar later 2 leerkrachten bereid deel te nemen aan de retentiestudie. Aldus namen 10 getrainde maar niet-gecoachte leerkrachten deel aan de retentiestudie. Van de 18 gecoachte leerkrachten uit de tweede evaluatiestudie namen 8 leerkrachten deel aan de retentiestudie. In totaal namen 18 getrainde leerkrachten (gecoacht of niet-gecoacht) aan de retentiestudie deel (zie Tabel 1). Deze leerkrachten gaven les in de groepen 3 tot en met 8 en waren afkomstig van 9 scholen.

De controlegroep bestond uit 11 leerkrachten die verbonden waren aan 7 scholen met vergelijkbare sociaal-economische en geografische achtergronden als de scholen uit de experimentele groepen (6 van de 9 leerkrachten uit de controlegroep van de eerste evaluatiestudie en 5 van de 14 leerkrachten uit de controlegroep van de tweede evaluatiestudie).

Ter controle op mogelijke zelf-selectie-effecten werden allereerst de scores op de voormeting betreffende het taakgerichte gedrag van de leerlingen en de instructie- en klasmanagementvaardigheden van de aan de retentiestudie deelnemende klassen en leerkrachten vergeleken met die van de niet aan de retentiestudie deelnemende klassen en leerkrachten. Uit de gegevens blijkt dat de scores van de experimentele groep die deelnam aan de retentiestudie ($n = 18$) op 5%-niveau niet significant afweken van de scores van de experimentele groep die niet deelnam aan de retentiestudie ($n = 27$). Voor de controlegroep werd slechts één significant verschil gevonden. De leerkrachten uit de controlegroep die deelnamen aan de retentiestudie ($n = 11$) scoorden significant lager dan de leerkrachten uit de controlegroep die niet deelnamen aan de retentiestudie ($n = 12$) op een van de vijf subschalen betreffende het management- en instructiegedrag, namelijk 'afstemming van de instructie'. Vervolgens werden de scores op de nameting van de aan de retentie-meting deelnemende en niet-deelnemende leerkrachten met elkaar vergeleken. Niet uitgesloten kon worden dat de aan de retentie-studie deelnemende leerkrachten zich onderscheiden door enthousiasme en inzet tij-

dens de training, wat zou kunnen resulteren in hogere scores op de nameting. Dit bleek niet het geval. Voor de nameting werden geen significante verschillen tussen de experimentele en de controlegroep aangetroffen. Op grond van deze vergelijkingen op de voor- en name-ting werd geconcludeerd dat er nauwelijks sprake was van zelf-selectie op de voor het onderzoek relevante variabelen.

Het aantal leerkrachten dat in de retentie-studie bereid was om schooltoetsen af te nemen, was groter dan het aantal leerkrachten dat bereid was om observatoren in de klas toe te laten. Van de 12 klassen in de retentiestudie met getrainde maar niet-gecoachte leerkrachten, hadden 7 leerkrachten aan de eerste evaluatie-studie deelgenomen en 5 aan de tweede evalua-tiestudie. Alle klassen in de retentiestudie met gecoachte leerkrachten ($n = 14$) namen ook deel aan de tweede evaluatiestudie (in de eerste evaluatiestudie vormde coaching geen experi-mentele conditie). De controlegroep bestond uit 14 leerkrachten (7 uit de eerste en 7 uit de tweede evaluatiestudie). Zie voor een overzicht Tabel 1.

Hoewel alle leerkrachten die deelnamen aan het observatie-onderzoek uit de eerste en de tweede evaluatiestudie benaderd werden, was het aantal leerkrachten dat bereid was tot deel-name aan de retentiestudie niet indrukwekkend (40% van de totale experimentele groep en 48% van de totale controlegroep uit beide eval-uatiestudies). Zoals reeds eerder opgemerkt kwam dit doordat de leerkrachten en scholen geen onmiddellijk profijt hadden van de reten-tiestudie. Andere argumenten om van deel-name af te zien waren: andere prioriteiten, overbelasting, ziekte en leerkrachten die de school hadden verlaten.

6.2 Instrumentarium

De meetinstrumenten die gebruikt werden om het effect van de nascholing of training vast te stellen op het taakgerichte gedrag van de leer-lingen en het instructie- en klasmanagement-gedrag van de leerkrachten waren identiek aan die van de eerste en tweede evaluatiestudie. Gebruikt werden een observatie-categorieën-systeem en een beoordelingsschaal. Om het ef-fect op de leerlingprestaties vast te stellen werd gebruik gemaakt van gestandaardiseerde toet-sen.

6.2.1 Het observatie-instrument

Voor de registratie van het taakgerichte gedrag van de leerlingen en de instructie- en klassema-nagementvaardigheden van de leerkracht werd gebruik gemaakt van het observatie-instrument COMMIT. Hiermee werd een schatting verkren-gen van de tijd die door de leerlingen en leer-krachten aan verschillende activiteiten en settings wordt besteed. De observatoren regi-streerden: 1) de taakgerichte leertijd van leer-lingen, 2) de groeperingsvorm (setting) waarin de leerlingen werkten, 3) de activiteit van de leerkracht en 4) de doelgroep van de leerkracht. Gebruikt werd een specifieke vorm van 'time-sampling', namelijk de 'predominant activity sampling'-procedure (Tyler, 1979). De pro-cedure verliep als volgt: eerst werden de activiteiten van leerling één en de leerkracht geobserveerd (duur: 7 seconden), waarna de geobserveerde gedragingen werden gecodeerd in de vier boven beschreven rubrieken van het observatie-instrument (duur: 13 seconden). Vervolgens werden de activiteiten van leerling twee en de leerkracht geobserveerd en geco-deerd, waarna de procedure werd herhaald tot elke leerling uit de klas was geobserveerd en gecodeerd. Hierbij werden de leerlingen in een vooraf vastgestelde volgorde geobserveerd.

Voorafgaande aan de observaties volgden drie observatoren een observatie-training van ongeveer 40 uur. Bij de training werd gebruik gemaakt van video-fragmenten en lesobservaties in reële klassituaties. De training werd ver-zorgd door een ervaren trainer uit de tweede evaluatiestudie die op zijn beurt getraind was door een ervaren trainer uit de eerste evaluatie-studie. Dit ter waarborging van een identieke observatie-procedure in de drie studies. De tussen-observator-betrouwbaarheid voor de verschillende observatie-categorieën, geschat met behulp van variantie-analyse (Winer, 1971), varieerde van .70 tot 1.00 (mediaan .97), met uitzondering van de categorie 'doelgroep hele klas' (.25). Deze categorie werd niet in de verdere analyse betrokken.

6.2.2 De beoordelingsschaal

Met behulp van de Management- en Instructie Schaal (MIS) werden meer specifieke gege-vens verzameld over de kwaliteit van de management- en instructievaardigheden van de leerkracht. De MIS was gebaseerd op onder-

zoek naar effectief managementgedrag van leerkrachten (Evertson, Emmer, Sanford & Clements, 1983; Veenman, Lem, Voeten, Winkelmoen & Lassche, 1986) en op onderzoek naar kenmerken van effectieve instructie (Rosenshine & Stevens, 1986; Good, Grouws & Ebmeier, 1983). De 31 items van de schaal representeerden vijf subschalen: 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. Deze items werden ook gebruikt in de eerste en tweede evaluatiestudie. De alpha-coëfficiënten voor deze subschalen varieerden van .66 tot .91 (zie Tabel 3). De tussen-observator-betrouwbaarheid varieerde van .63 tot .81 (mediaan .74).

6.2.3 Schooltoetsen

Om het effect van de nascholing op de leerprestaties vast te stellen werden in de groepen 3 tot en met 8 drie toetsen afgenomen: een voor technisch lezen (de Een-Minuut-Test (EMT) van Brus en Voeten), een voor begrijpend lezen (CITO-Begrijpend lezen, Begrippentest Aarnoutse, Begrijpend lezen Aarnoutse en Hoofdgedachtetest Aarnoutse) en een voor rekenen (Tempo-Test-Rekenen van De Vos). Voor alle groepen was de EMT de enige identieke test. Van de Tempo-Test-Rekenen werd de gehele toets afgenomen in de groepen 5 tot en met 8, in de groepen 3 en 4 een gedeelte van de test. Voor begrijpend lezen bestonden geen toetsen die voor alle groepen van de basisschool geschikt waren. Hier werden verschillende toetsen gebruikt; elke test trachtte uitgaande van een bepaald leerjaar de vaardigheid van begrijpend lezen te meten.

6.3 Dataverzameling en -verwerking

De voormeting van de eerste evaluatiestudie vond plaats in de maanden november-december 1986, de nameting in de maanden mei - juni 1987. Gedurende de tweede evaluatiestudie vond de voormeting plaats in de maanden november-december 1989 en de nameting in de maanden mei-juni 1990. De meting voor de retentiestudie vond plaats in de periode mei-juli 1992, vijf jaar na de nameting van de eerste evaluatiestudie en twee jaar na de nameting van de tweede evaluatiestudie.

Alle leerkrachten die deelnamen aan de re-

tentiestudie werden geobserveerd tijdens een reken- en een lees/taalles. Alle observatiegegevens die werden verkregen met behulp van de COMMIT hadden per meting betrekking op de gemiddelde tijdsbesteding per klas of per leerkracht, berekend over de twee geobserveerde lessen. Per observatiecategorie resulteerde een tijdschatting die werd uitgedrukt in minuten (geobserveerd werden lessen van 40 minuten). De observatiegegevens van de MIS werden per subschaal omgezet in somscores (berekend door de itemscores voor elk van de vijf subschalen te combineren). Deze scores hadden eveneens betrekking op de gemiddelden van twee lessen.

Bij de toetsing van verschillen tussen de experimentele groep en de controlegroep en tussen de experimentele groepen onderling werd een significantieniveau van 5% aangehouden (eenzijdig). Voor de observatiegegevens vormde de klas of leerkracht de analyse-eenheid; voor de gegevens van de schooltoetsen werd de leerling als analyse-eenheid genomen. Voor de toetsing van verschillen werd gebruik gemaakt van het programma MANOVA van SPSS-X. Voor meer gedetailleerde gegevens over de opzet, het instrumentarium en de dataverzameling wordt verwezen naar Raemaekers en Veenman (1994).

7 Resultaten

Ter toetsing van het effect van het nascholingsprogramma op de taakgerichte leertijd van de leerlingen en het instructie- en klasmanagementgedrag van de leerkracht werden covariantie-analyses (ANCOVA) uitgevoerd met de scores op de voormeting als covariabelen. De scores van de nameting en de retentiemeting vormden de afhankelijke variabelen, met 'treatment' (training of geen training) als onafhankelijke variabele. Uitgevoerd werden drie vergelijkingen: (1) experimentele groep versus controlegroep, (2) coaching versus geen coaching en (3) de nascholingsgroep van 1986/87 versus de nascholingsgroep van 1989/90. Opgemerkt moet worden dat alleen de scores van de niet-gecoachte leerkrachten uit de tweede evaluatiestudie vergeleken konden worden met die van de leerkrachten uit de eerste evaluatie-

studie, omdat coaching geen experimentele conditie vormde in de eerste evaluatiestudie.

Allereerst werd getoetst of er op de voormeting verschillen bestonden tussen de experimentele groep (getrainde leerkrachten met of zonder coaching) en de controlegroep. Er werden geen significante aanvangsverschillen gevonden tussen beide groepen. Tussen de experimentele groep met en zonder coaching werd op de voormeting alleen voor de categorie 'procedurele activiteiten' een significant verschil aangetroffen. Bij de gecoachte leerkrachten besteden de leerlingen tijdens de voormeting significant minder tijd aan procedurele activiteiten dan leerlingen bij niet-gecoachte leerkrachten. Verder scoorden gecoachte leerkrachten op de voormeting significant hoger op een van de vijf subschalen van de MIS dan niet-gecoachte leerkrachten, namelijk op de sub-schaal 'gebruik van materiaal en ruimte'. Uit deze resultaten van de toetsing op aanvankelijke verschillen blijkt dat de controle- en de experimentele groep(en) op de voor het onderzoek relevante variabelen nauwelijks van elkaar verschilden. Enkele kleine verschillen werden aangetroffen tussen gecoachte en niet-gecoachte leerkrachten.

Ter beantwoording van de vraagstellingen 1-3 werden eerst vergelijkingen uitgevoerd tussen de onderscheiden groepen voor de scores op de nametingen (nameting én retentiemeting). Vervolgens werden de verschillen getoetst tussen nameting en retentiemeting om te zien of eventuele effecten van de nascholing een duurzaam karakter hadden of niet.

7.1 Effecten van de nascholing

Tabel 2 bevat een overzicht van de gegevens betreffende de tijdsbesteding voor de voor het onderzoek relevante variabelen van het observatie-instrument COMMIT en de beoordelingsschaal MIS. Tabel 3 geeft de resultaten weer van de uitkomsten van de statistische toetsen (uitgedrukt in *t*-waarden). De scores van de nameting en de retentie-meting werden gemiddeld tot één score ter bepaling van het algemeen effect van de nascholing. Getoetst werden de verschillen in deze scores tussen de experimentele groep en de controlegroep en tussen de experimentele groepen onderling, waarbij de scores van de voormeting fungeerden als covariaat. De uitkomsten van deze toet-

singen staan vermeld onder het hoofd 'trainingseffecten'. Bij de interpretatie van deze uitkomsten dienen de gegevens uit Tabel 2 betrokken te worden.

Effecten van het nascholingsprogramma

De gegevens betreffende de vergelijking tussen de experimentele groep (gecoachte en niet-gecoachte leerkrachten) en de controlegroep laten zien dat het nascholingsprogramma een significant effect heeft op het taakgerichte gedrag van de leerlingen. De leerlingen in de experimentele groep zijn meer taakgericht bezig dan de leerlingen in de controlegroep: 80% (of 32.0 minuten: de gemiddelde score van de na-en retentiemeting) versus 72% (28.7 minuten). Dit verschil is significant ($p < .01$). Leerlingen in de experimentele groep besteden significant minder tijd aan wachten op de leerkracht en aan niet-taakrelevante activiteiten.

In Tabel 3 worden ook de uitkomsten weergegeven met betrekking tot de gemiddelde taakgerichte leertijd in de twee onderscheiden settings: instructie/uitleg en zelfstandige verwerking. In beide settings zijn de leerlingen in de experimentele groep significant taakgerichter dan leerlingen in de controlegroep. Gedurende de uitleg-fase van de les besteden leerlingen in de experimentele groep gemiddeld 86% van de geobserveerde tijd aan hun leertaken (34.4 minuten) en gedurende de zelfstandige verwerkingsfase van de les gemiddeld 78% (31.0 minuten). Voor de leerlingen uit de controlegroep bedragen deze cijfers respectievelijk 77% (30.9 minuten) en 66% (26.6 minuten). Deze resultaten laten zien dat de leerlingen van leerkrachten die aan de nascholing hebben deelgenomen meer taakgericht zijn dan de leerlingen van leerkrachten die niet aan de nascholing hebben deelgenomen.

Gegevens over aspecten van het instructie- en klasmanagementgedrag van leerkrachten zijn eveneens opgenomen in Tabel 3. Significante verschillen zijn aangetroffen tussen de experimentele en de controlegroep voor de categorieën 'terugblik' en 'niet-actief lesgedrag'. De leerkrachten uit de experimentele groep besteden significant meer tijd aan het houden van een terugblik en significant minder tijd aan niet-op-instructiegerichte activiteiten dan de leerkrachten uit de controlegroep (5% versus 3% en 1% versus 6%).

Tot slot bevat Tabel 3 de resultaten van de Management- en Instructie Schaal (MIS). Op vier van de vijf subschalen behalen getrainde leerkrachten significant hogere scores dan niet-getrainde leerkrachten. Alleen voor de subschaal 'gebruik van materiaal en ruimte' is geen significant verschil gevonden. Geconcludeerd kan worden dat de training een positief effect heeft gehad op de instructie- en de klasmanagementvaardigheden van de leerkrachten.

Effecten van coaching

Tussen gecoachte en niet-gecoachte leerkrachten zijn geen significante verschillen aangetroffen. Coaching heeft geen effect gehad op de taakgerichte leertijd van de leerlingen of op het instructie- en klasmanagementgedrag van de aan de nascholing deelnemende leerkrachten.

Effecten van de toepassingstijd

Aangenomen werd dat leerkrachten die vijf jaar geleden aan de nascholing hadden deelgenomen door een langere toepassingstijd hoger zouden scoren dan leerkrachten die twee jaar geleden aan de nascholing hadden deelgenomen. Deze verwachting is niet bewaarheid. Tussen de beide groepen van leerkrachten zijn geen significante verschillen gevonden. Hierbij moet echter worden aangetekend dat de experimentele groep van 1989/90 slechts uit twee leerkrachten bestaat. De uitkomsten van deze vergelijking dienen derhalve met de nodige voorzichtigheid te worden geïnterpreteerd.

7.2 Retentie-effecten

Om de duurzame effecten van het nascholingsprogramma vast te stellen werden de verschillen tussen de retentie-scores en de scores van de nameting (uitgedrukt als verschillen) van de experimentele groep (gecoachte en niet-gecoachte leerkrachten) en de controlegroep met elkaar vergeleken, eveneens de verschillen tussen de experimentele groepen onderling. Hierbij werden steeds de scores van de voormeting gebruikt als covariaten. De uitkomsten van deze vergelijkingen zijn samengevat in Tabel 3 onder het hoofd 'retentie-effecten'.

Effecten van het nascholingsprogramma op de retentie

Na controle van mogelijke verschillen op de

voormeting, zijn tussen de experimentele groep en de controlegroep geen significante verschillen aangetroffen. Men mag hieruit afleiden dat de scores van de getrainde leerkrachten in de periode tussen de nameting en de retentiemeting, in vergelijking met die van niet-getrainde leerkrachten, niet significant zijn toe- of afgenomen. De resultaten die behaald zijn direct na de training blijven voor de meeste variabelen na verloop van tijd op hetzelfde niveau liggen.

Effecten van coaching op de retentie

Tussen gecoachte en niet-gecoachte leerkrachten zijn geen significante verschillen gevonden. Ondanks de toevoeging van coaching aan de nascholing blijken gecoachte leerkrachten tussen nameting en retentiemeting niet meer vooruit of achteruit te gaan dan niet-gecoachte leerkrachten.

Effecten van de toepassingstijd op de retentie

Drie significante verschillen zijn geconstateerd tussen de getrainde leerkrachten uit 1986/87 en 1989/90. De getrainde leerkrachten uit 1989/90 scoorden lager in de categorie 'wachten op de leerkracht' en hoger in de categorie 'geleide oefening' en op de subschaal 'instructievaardigheden' van de MIS. De gevonden verschillen zijn tegengesteld aan de geformuleerde verwachting. Leerkrachten die de nascholing twee jaar geleden hadden gevolgd scoorden op deze variabelen tussen de nameting en de retentiemeting hoger dan de leerkrachten die de nascholing vijf jaar geleden hadden gevolgd. Gezien het aantal variabelen dat in de toetsing is betrokken, zijn de verschillen tussen beide groepen niet groot.

7.3 Transferpatronen

In Figuur 1 zijn enkele effecten van de training schematisch weergegeven door middel van transferpatronen. Deze transferpatronen maken zichtbaar welke veranderingen tussen voor-, na- en retentiemeting zijn opgetreden. Het transferpatroon met betrekking tot de taakgerichte leertijd van de leerlingen en het geven van een terugblik door de leerkracht kan in termen van Den Ouden (1992) omschreven worden als een 'veranderingscontinueringspatroon' en wel als een 'voortgezet veranderingsproces'. Bij 'geleide oefening' is sprake van

een gestaakt veranderingsproces: direct na de training vindt verandering van gedrag plaats maar deze verandering verdwijnt na verloop van tijd. Er treedt geen transfer op. Bij 'gerichte controle' ook niet. Hier vindt geen implementatie en dus ook geen transfer plaats. Er is geen intentie tot verandering te constateren. De twee weergegeven gedragspatronen voor 'instructievaardigheden' en de 'organisatie van de instructie' zijn indicatief voor de scores op de Management- en Instructie Schaal (MIS). Op alle vijf subschalen is sprake van een 'veranderingscontinueringspatroon'. Alleen bij het 'gebruik van materiaal en ruimte' laat de controlegroep tussen de nameting en de retentiemeting een sterkere stijging zien op deze subschaal van de MIS dan de experimentele groep (figuur hier niet opgenomen; zie Tabel 2).

7.4 Leerprestaties

Het effect van het nascholingsprogramma op de leerprestaties van de leerlingen (vraagstelling 4) is samengevat in Tabel 4. Om verschillen in testlengte en testsoorten voor rekenen en begrijpend lezen te verdisconteren, werden alle ruwe leerlingsscores gestandaardiseerd tot z-scores. Verschillen tussen de klassen werden getoetst met behulp van variantie-analyse (eenzijdig).

Bij middeling van de retentie-scores over de klassen heen zijn geen significante verschillen gevonden. De klassen met getrainde leerkrachten scoren niet hoger op de drie schooltoetsen dan de klassen met niet-getrainde leerkrachten. Ook per klas of groep zijn geen systematische verschillen in retentie-scores gevonden tussen getrainde en niet-getrainde leerkrachten (zie Raemaekers & Veenman, 1994). In eerder onderzoek zijn geen verschillen geconstateerd in de scores voor de nameting (Roelofs, Veenman & Voeten, 1994). Verder zijn in deze retentiestudie geen verschillen in leerprestaties gevonden tussen de klassen met gecoachte of niet-gecoachte leerkrachten en tussen de klassen uit de experimentele groep van 1986/87 en de experimentele groep van 1989/90. Bij de laatste vergelijking zijn alleen niet-gecoachte leerkrachten betrokken, omdat coaching in de experimentele groep van 1986/87 niet werd toegepast. Deze resultaten met betrekking tot de leerprestaties van de leerlingen dienen met de nodige voorzichtigheid te worden geïnterpre-

teerd omdat gegevens betreffende leerprestaties voor de aanvang van de training niet voorhanden zijn. Om deze reden heeft geen controle kunnen plaatsvinden op aanvankelijke verschillen in leerprestaties tussen de groepen².

8 Discussie

In deze studie werden de lange-termijn effecten van het nascholingsprogramma onderzocht. Aan de hand van de verzamelde observatiegegevens - twee en vijf jaar na afloop van de nascholing - kon worden vastgesteld dat de getrainde leerkrachten nog steeds gebruik maakten van de in de nascholing geleerde instructie- en klasmanagementvaardigheden. De meeste van de tijdens de nascholing geleerde vaardigheden bleken een duurzame toepassing gevonden te hebben in de klas.

Eén strategie voor transferbevordering bleek echter niet succesvol. Het geven van coaching aan leerkrachten resulteerde niet in duurzame effecten op het taakgerichte gedrag van de leerlingen of op de onderwijsvaardigheden van de leerkrachten. Deze uitkomst kan mogelijk verklaard worden doordat de leerkrachten na afloop van de nascholing geen coaching meer ontvingen. Het ontbrak de schoolbegeleiders aan mogelijkheden voor follow-up coaching. Blijvende ondersteuning op de werkplek kon niet gerealiseerd worden, waardoor de direct na de training gevonden coachingseffecten na verloop van tijd 'verdampen.'

Effecten als gevolg van de toepassingstijd werden nauwelijks gevonden. Waar verschillen werden gevonden waren deze in het voordeel van de groep van getrainde leerkrachten die het nascholingsprogramma twee jaar geleden hadden gevolgd. Een mogelijke verklaring kan zijn dat de leerkrachten uit 1989/90 het belang van de betreffende vaardigheden nog niet vergeten waren. Gezien de omstandigheid dat deze groep slechts twee leerkrachten telde, dient een hypothese omtrent verdere perfectionering van getrainde vaardigheden in de tijd in toekomstig onderzoek nader getoetst te worden.

Vergeleken met de leerkrachten uit de controlegroep boekten de leerkrachten uit de experimentele groep tussen de nameting en de retentiemeting geen verdere vooruitgang met

betrekking tot de uitvoering van de gewenste instructie- en klasmanagementvaardigheden. Uit de resultaten met betrekking tot het interview-instrument 'Gebruiksniveaus van innovatie' (Hall & Loucks, 1977; Hord, Rutherford, Huling-Austin & Hall, 1987) dat in de reëntiestudie eveneens werd gebruikt, bleken de getrainde leerkrachten te functioneren op niveau 3 (mechanisch gebruik) en niveau 4A (routine) (zie Raemaekers & Veenman, 1994). Een stabiel, routinematig patroon van gebruik bleek de werkwijze van de leerkrachten het beste te karakteriseren. Dit niveau van gebruik kan mogelijk verklaard worden door de wijze waarop de instructie- en klasmanagementvaardigheden werden geoperationaliseerd. De meeste vaardigheden waren zeer concreet omschreven. Dergelijke nauw omschreven vaardigheden in gedragstermen bieden mogelijk weinig ruimte voor verdere verfijning of voor een frequenter toepassing. In een les, bijvoorbeeld, is gerichte controle afhankelijk van de problemen die leerlingen hebben met zelfstandig werken. Geen enkele leerkracht zal zijn leerlingen minder zelfstandig laten werken om veel gerichte controle te kunnen toepassen.

Hoewel de getrainde leerkrachten in hun klassen een hogere taakgerichte leertijd realiseerden dan niet-getrainde leerkrachten, leidde dit verschil niet tot een verschil in leerprestaties. Dat meer taakgerichte leertijd niet resulteerde in meer leerwinst kan mogelijk op twee manieren verklaard worden. De eerste verklaring is dat de leerkrachten méér leertijd als een doel op zich beschouwd hebben, wat geleid kan hebben tot meer van hetzelfde. In deze studie zijn geen gegevens verzameld over de wijze waarop de taakgerichte leertijd werd ingevuld. We weten derhalve niet of de geboekte winst aan leertijd besteed werd aan de juiste leertaken. Leertijd werd op een kwantitatieve en niet op een kwalitatieve wijze gemeten. In toekomstige nascholingsprogramma's dienen de leerkrachten meer aandacht te besteden aan de vraag: tijd voor welke leertaken? De kwaliteit van de uitgevoerde leertaak bepaalt wat en hoeveel er geleerd wordt. De tweede verklaring is dat de nascholing het primaat legde bij de instructie van de leerkracht en niet bij het leren van de leerlingen. Leerkrachten zouden meer moeten denken in termen van de gewenste denk- of leerprocessen die ze bij de leerlingen

willen realiseren. Voor toekomstige trainingen betekent dit een verschuiving van denken in instructie-activiteiten naar denken in leeractiviteiten (cf. Simons, 1995). Op deze wijze wordt de invulling van de taakgerichte leertijd direct gekoppeld aan het leren van de leerlingen.

Een van de vaardigheden waarvan het belang voor combinatieklassen wordt benadrukt is het geven van geleide oefening. Twee tot vijf jaar na de training blijkt het effect van de nascholing voor deze vaardigheid verdwenen te zijn (zie Figuur 1). Geleide oefening is belangrijk voor het geven van onderwijs in combinatieklassen omdat via geleide oefening de leerlingen worden voorbereid op het zelfstandig verwerken of toepassen van de leerstof. Pas wanneer de leerkracht zich er van overtuigd heeft dat de leerlingen de leerstof begrepen hebben, kan hij de leerlingen in een groep zelfstandig laten werken en zich wenden tot een andere groep. Uit de gegevens van de voormeting blijkt (zie Tabel 2) dat de leerkrachten weinig gebruik maakten van geleide oefening. Direct na de nascholing werd er significant meer gebruik van gemaakt, maar vergeleken met de voormeting niet overtuigend meer. Twee tot vijf jaar na de nascholing is dit effect verdwenen. Deze terugval in gebruik kan mogelijk verklaard worden doordat de leerkrachten weinig succesvol zijn geweest bij de toepassing van deze vaardigheid in de complexe organisatie van een klas met twee of meer groepen. Het kan ook zijn dat de leerkrachten te weinig ondersteuning uit de werkomgeving hebben ontvangen. In de opzet en uitvoering van de nascholing is de meeste aandacht uitgegaan naar het nemen van transferbevorderende maatregelen vóór en tijdens de nascholing. Aan het traject is weinig aandacht besteed. Zo hebben leerkrachten geen follow-up training ontvangen, is weinig gedaan aan nazorg. In de toekomst zou daarom meer aandacht gegeven dienen te worden aan transferbevorderende organisatiemaatregelen, maatregelen die liggen op het niveau van de school als organisatie (zoals follow-up activiteiten, evaluatie, feedback, peer-coaching).

Noten

- 1 De eerste twee versies van het trainingsprogramma verschenen onder de titel *Omgaan met Combinatieklassen*. Na de twee evaluatiestudies werd het programma herzien en werd het vanaf de derde druk uitgebracht onder de titel *Effectieve instructie en doelmatig klasmanagement* (Veenman, Lem, Roelofs & Nijssen, 1994).
- 2 In de retentiestudie konden de scores voor de schooltoetsen uit de tweede evaluatiestudie niet als covariaten gebruikt worden omdat de meeste leerkrachten twee en vijf jaar na afloop van de nascholing een andere klas hadden met andere leerlingen.

Literatuur

- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research, *Personnel Psychology*, 41(1), 63-105.
- Bennett, B. (1987). *The effectiveness of staff development training practices: A meta-analysis*. Unpublished doctoral dissertation, University of Oregon.
- Broad, M. L., & Newstrom, J. W. (1992). *Transfer of training: Action-packed strategies to ensure high payoff from training investments*. Reading: Addison-Wesley.
- Commissie Evaluatie Basisonderwijs. (1994). *Inhoud en opbrengsten van het basisonderwijs*. Leiden: Distributiecentrum DOP.
- Evertson, C. M., Emmer, E. T., Sanford, J. P., & Clements, B. S. (1983). Improving classroom management: An experiment in elementary school classrooms. *Elementary School Journal*, 84, 173-188.
- Good, T. L., Grouws, D. A., & Ebmeier, H. (1983). *Active mathematics teaching*. New York: Longman.
- Hall, G. E., & Loucks, S. F. (1977). A developmental model for determining whether the treatment is actually implemented. *American Educational Research Journal*, 14, 263-276.
- Hord, S. M., Rutherford, W. L., Huling-Austin, L., & Hall, G. E. (1987). *Taking charge of change*. Alexandria: Association for Supervision and Curriculum Development.
- Joyce, B. R., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership*, 37, 379-385.
- Joyce, B. R., & Showers, B. (1995). *Student achievement through staff development: Fundamentals of school renewal* (2nd ed.). New York: Longman.
- Ouden, M. D. den (1992). *Transfer na bedrijfsopleidingen: Een veldonderzoek naar de rol van voornemens, sociale norm, beheersing en sociale steun bij opleidingstransfer*. Amsterdam: Thesis Publishers.
- Raemaekers, J., & Veenman, S. (1994). *Institutionalisering van teamgerichte nascholing: Duurzame effecten van de cursus Omgaan met Combinatieklassen*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Roelofs, E., Veenman, S., & Raemaekers, J. (1993). Teamgerichte nascholing en coaching: Effecten op het didactisch handelen van leerkrachten in combinatieklassen. *Pedagogische Studiën*, 70, 352-367.
- Roelofs, E., Veenman, S., & Voeten, M. (1994). Effecten van een teamgerichte nascholing en van coaching op leerprestaties. *Pedagogische Studiën*, 71, 442-456.
- Rosenshine, B. V., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376-391). New York: MacMillan.
- Simons, P. R. J. (1992). De toepassing van transferbevorderende maatregelen in arbeidsorganisaties. *Tijdschrift voor Onderwijsresearch*, 17, 26-37.
- Simons, R. J. (1995). *Leren en werken*. Nijmegen: Vakgroep Onderwijskunde, K.U. Nijmegen.
- Tulder, M. van. (1992). *Nascholing en onderwijsvernieuwing*. (Academisch proefschrift, K.U. Nijmegen) Nijmegen: Universiteitsdrukkerij.
- Tyler, S. (1979). Time-sampling: A matter of convention. *Animal Behavior*, 27, 801-810.
- Veenman, S., Lem, P., & Roelofs, E. (1990). Omgaan met combinatieklassen: Een onderzoek naar de effecten van een teamgericht nascholingsprogramma. *Pedagogische Studiën*, 67, 45-58.
- Veenman, S., Lem, P., Roelofs, E., & Nijssen, F. (1994). *Effectieve instructie en doelmatig klasmanagement: Een schoolverbeteringsprogramma voor eenvoudige en combinatieklassen* (5de druk). Amsterdam: Swets & Zeitlinger.
- Veenman, S., Lem, P., Voeten, M., Winkelmolen, B., & Lassche, H. (1986). *Onderwijs in combinatieklassen*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs.
- Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

Manuscript aanvaard 22-9-1995

Auteurs

S. Veenman is als universitair hoofddocent werkzaam bij de Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen.

J. Raemaekers was ten tijde van dit onderzoek verbonden als onderzoeksmedewerker aan de Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen. Momenteel is hij als schoolbegeleider werkzaam bij de Stichting Schoolpedagogisch Centrum Westelijke Mijnstreek te Sittard.

Adres: Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen, Postbus 9104, 6500 HE Nijmegen.

Abstract

Retention effects of a staff-development program for teachers in multi-grade classes.

S. Veenman & J. Raemaekers. *Pedagogische Studiën*, 1996, 73, 357-371.

This study describes the long-term effects of a staff-development program based on selected findings

from teaching-effectiveness research in schools with multi-grade or mixed-age classes. The short-term effects of this program were examined in two studies directed at schools with multi-grade classes. The first improvement study was conducted in the school year 1986/87; the second improvement study was conducted in 1989/90. In the latter study, the effects of coaching in addition to participation in the staff-development program were also evaluated. In 1992, a retention or follow-up study was conducted. A quasi-experimental, treatment-control group design was used to test the long-term effects of the program *Dealing with Multi-Grade Classes* and the effects of coaching. Based on pre-, post-training classroom observations, the follow-up study revealed a significant treatment effect for the time-on-task levels of the pupils in the multi-grade classes and for the instructional and classroom management skills of the teachers. No significant differences were found between the coached and uncoached teachers and between the teachers who followed the program either two or five years ago. No significant differences were found between the post-test and the retention test. This suggests that the training results were quite stable. No significant differences in achievement were found between the pupils in classes with trained teachers and the pupils in classes with untrained teachers.

Appendix

Tabel 1

Overzicht van het meetinstrumentarium en het aantal deelnemende leerkrachten aan de eerste en tweede evaluatiestudie en de retentiestudie

Instrument	Eerste studie (1986/87)			Tweede studie (1989/90)			Retentiestudie (1992)		
	Exp. 1 <i>n</i>	Exp. 2 <i>n</i>	Con. <i>n</i>	Exp. 1 <i>n</i>	Exp. 2 <i>n</i>	Con. <i>n</i>	Exp. 1 <i>n</i>	Exp. 2 <i>n</i>	Con. <i>n</i>
COMMIT (observatie)	17	nvt	9	10	18	14	10	8	11
MIS (observatie)	17	nvt	9	10	18	14	10	8	11
Toetsen (leerprestaties)	nvt	nvt	nvt	23	24	16	12	14	14

Opmerking: Exp. 1 = experimentele groep met niet-gecoachte leerkrachten; Exp. 2 = experimentele groep met gecochte leerkrachten; Con. = controlegroep

Gemiddelde tijdsbesteding voor de categorieën van de COMMIT en de subschalen van de MIS voor de voormeting, nameting en retentiemeting

Opmerking: Lesduur = 40 minuten. Voor = voormeting, Na = nameting, Ret = retentiemeting

Tabel 3

Training- en retentie-effecten: Resultaten van de t-toetsen voor de scores van de observatiecategorien van de COMMIT en van de subschalen van de Management en Instructie Schaal (MIS) van de retentie-studie (met de score van de voormeting als covariaat)

Observatiecategorien	Trainingseffecten			Retentie-effecten		
	Training: exp. vs. con.	Coaching: gecoacht vs niet- gecoacht	Tijd: 1986/87 vs. 1989/90	Training: exp. vs.con.	Coaching: gecoacht vs.niet- gecoacht	Tijd: 1986/87 vs. 1989/90
	t	t	t	t	t	t
COMMIT						
<i>Leerlinggedrag</i>						
Taakgericht	3.83**	-0.28	-0.19	-1.29	0.69	-1.60
Procedurale activiteiten	-1.69	0.24	-0.10	1.15	0.38	-0.32
Wachten op de leerkracht	-2.19*	-0.20	0.12	-0.11	-0.31	2.08*
Niet met de opdracht bezig	-3.14**	0.26	-0.01	1.07	-1.19	1.84
<i>Setting en taakgerichtheid</i>						
Instructie en taakgericht (%)	3.44**	-0.36	0.41	-0.80	-0.14	-0.52
Verwerking en taakgericht (%)	3.86**	-0.37	0.39	-1.97	1.02	-1.44
<i>Leerkrachtgedrag</i>						
Terugblik	4.93**	1.78	-2.05	-0.26	1.65	-0.60
Geleide oefening	1.56	-1.75	0.68	-1.47	0.91	-2.56*
Gerichte controle	-0.25	0.54	0.43	-0.45	0.96	1.22
Organisatie/leswisseling	-1.52	1.11	0.89	-0.09	0.37	0.27
Niet-actief lesgedrag	-2.59*	0.27	-0.09	1.52	0.29	1.28
<i>Subschalen MIS</i>						
Instructievaardigheden ($\alpha = .90$)	4.89**	-0.88	-0.87	-0.27	1.04	-3.75**
Organiseren vd instructie ($\alpha = .84$)	4.10**	0.22	-0.03	0.56	0.73	-0.71
Gebruik materiaal/ruimte ($\alpha = .66$)	1.95	1.00	-0.24	-0.83	-0.46	-0.62
Afstemming vd instructie ($\alpha = .86$)	3.37**	0.24	-0.22	-0.48	0.69	-0.91
Omgaan met storingen ($\alpha = .91$)	3.12**	0.09	-0.90	-0.51	0.78	-1.62

Opm.: Exp. = experimentele groep; Con. = controlegroep. Tijd = toepassingstijd.

De leerkrachtgedragingen uitleg en individuele hulp vormden qua tijdsbesteding geen oogmerk van de cursus en zijn derhalve niet in de toetsing opgenomen.

* = $p < .05$; ** = $p < .01$

Tabel 4

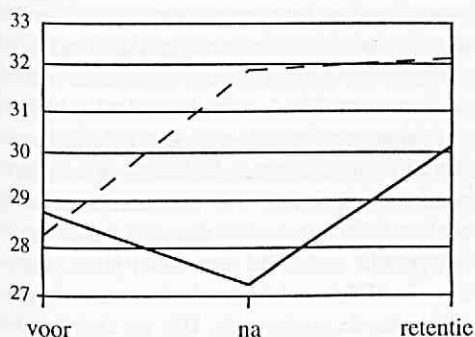
ANOVA-testresultaten voor de z-scores betreffende de leerprestaties (gemiddeld over de groepen drie tot en met acht)

Test	Experimentele groep			Controlegroep			F	p
	n	Gem	SD	n	Gem	SD		
Technisch lezen	630	0.01	1.00	349	-0.02	0.99	0.21	.64
Begrijpend lezen	634	0.04	1.01	329	-0.07	0.97	2.63	.11
Rekenen	632	-0.02	1.00	337	0.04	1.00	0.90	.34

Opm.: Experimentele groep n = 26 klassen, controlegroep n = 14 klassen.

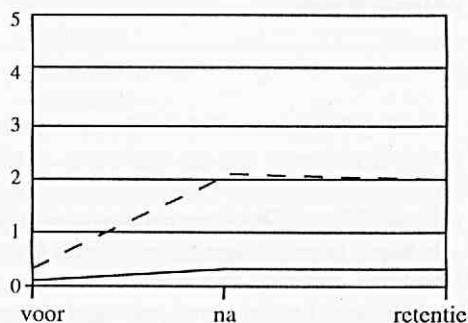
Taakgerichte leertijd

(lesduur 40 minuten)



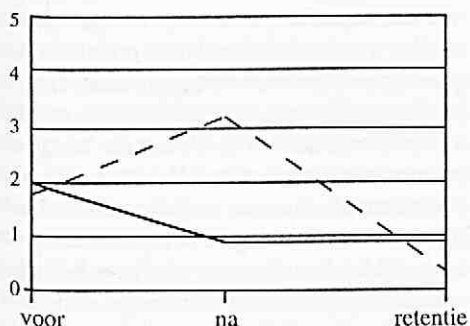
Terugblik

(lesduur 40 minuten)



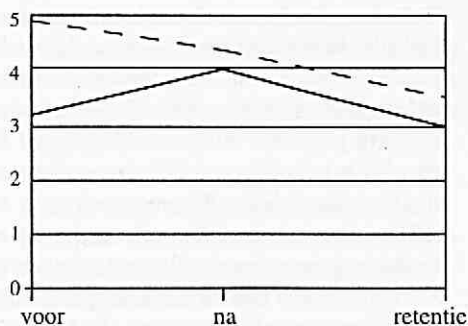
Geleide oefening

(lesduur 40 minuten)



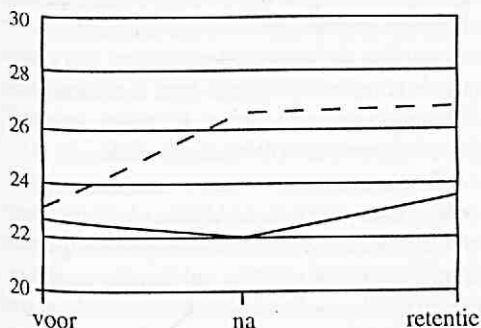
Gerichte controle

(lesduur 40 minuten)



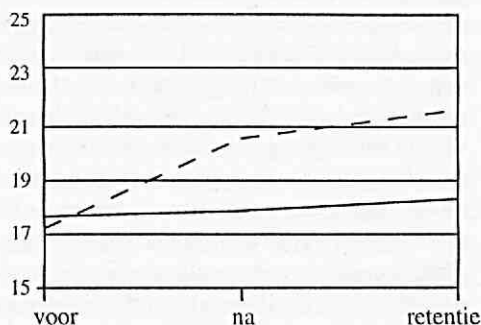
Instructievaardigheden (MIS)

(min.= 7, max.= 35)



Organisatie van de instructie (MIS)

(min.= 5, max.= 25)



--- Experimentele groep — Controlegroep

Figuur 1. Enkele voorbeelden van transferpatronen

Het AERA-Congres 1996, New York

Inleiding (N. Verloop, ICLON, Rijksuniversiteit Leiden)

Het jaarlijkse AERA-congres is zonder twijfel de meest grootschalige happening op het gebied van onderwijsonderzoek ter wereld. Al enkele jaren ligt het aantal presentaties tussen de vier- en de vijftienduizend. Het congres dat van 8 tot 12 april in New York werd gehouden vormde hierop, uiteraard, geen uitzondering. Opvallend is dat de Nederlandse onderwijsonderzoekers daarbij steeds sterk vertegenwoordigd zijn. Ik heb geen analyse gemaakt van het relatieve aandeel van de Nederlandse onderzoekers, maar wijs bijvoorbeeld op een analyse die vorig jaar door White in *Educational Researcher* is gepubliceerd. Hij ging onder meer na uit welke landen de presentaties op het AERA-congres van 1994 afkomstig waren. Omdat er geen reden is om aan te nemen dat deze getallen van jaar tot jaar ingrijpend zullen verschillen, vermeld ik hier kort zijn bevindingen. In totaal werden door auteurs uit 46 landen presentaties verzorgd. Het overgrote deel was vanzelfsprekend afkomstig uit de VS. In de totale lijst van 46 landen stond Nederland (met 37 presentaties) op de vijfde plaats, na de VS, Canada, Australië en Engeland. Met andere woorden, van alle niet-Engelssprekende landen scoorde Nederland het hoogste. Als we bovendien in aanmerking nemen dat het afwijzingspercentage voor veel divisies inmiddels al rond de 40% ligt en dat voorstellen van buiten de VS niet bepaald extra welwillend worden bejegend, kan dit als een opvallende prestatie gelden. Het geeft in elk geval weinig grond voor het pessimisme omtrent de internationale oriëntatie van de Nederlandse onderwijskunde, waarmee sommige beschouwers van het Nederlandse onderwijsonderzoek de laatste tijd behept lijken te zijn.

Het voortdurend speuren naar nieuwe formats die afwijken van de traditionele symposium-, paper- en posterpresentaties had ook

deze keer weer nieuwe varianten opgeleverd. Inmiddels komt men in het programmaboek zo'n 20 formats tegen, waaronder 'book sessions' (waarin een nieuwe publicatie wordt gepresenteerd en becommentarieerd), 'consultation sessions' (waarin de mogelijkheid voor het verkrijgen van hulp en advies gegeven wordt) en 'demonstrations'. Niet alle varianten blijken overigens erg zinvol te zijn. Dat gold wel voor een initiatief waarmee dit jaar voor het eerst werd geëxperimenteerd, nl. het als zodanig aankondigen van sessies die sterk op de praktijk gericht waren. Al een aantal jaren stimuleert de AERA praktiserende docenten tot deelname aan de conferentie. Dit, op zich zonder twijfel zinnige, streven tot vergroting van de functionaliteit van het onderwijsonderzoek leidt er soms toe dat onderzoekers die geïnteresseerd zijn in de onderzoeksmatige aanpak van een bepaalde thematiek, verzeild raken in sessies waarin uitsluitend over praktische toepassingsproblemen wordt gesproken. Dit jaar heeft men getracht door middel van speciale symbolen die sessies te labelen die 'of special interest to teachers' waren. Het systeem werkte weliswaar nog lang niet perfect, maar het biedt, bij striktere toepassing in de toekomst, wel een mogelijkheid om het type sessies te bezoeken waarin men geïnteresseerd is.

Het conferentiethema van dit jaar was "Research for education in a democratic society". Het kiezen van een thema voor een conferentie als deze is volstrekt zinloos. Omdat de divisies vrij zijn hun eigen beleid te voeren heeft een dergelijk thema geen enkele invloed op de voorstellen die aangeboden worden. Het enige wat de congresleiding kan doen is achteraf een label plakken op de sessies die 'theme-related' zijn, hetgeen dan ook braaf gebeurde.

Tot slot vermeld ik nog een elk jaar terugkerend hoogtepunt van de AERA, nl. de tentoonstelling waarop werkelijk alle grote uitgeverijen hun, veelal zeer recente, publicaties presenteren en verkopen. Dat laatste leidt er stevast toe dat tal van Europese congresgangers met aanzienlijk zwaardere koffers huiswaarts keren dan waarmee zij gekomen zijn.

In de hier volgende rapportage is gekozen voor een divisie-gewijze aanpak in plaats van een aanpak waarbij over allerlei specifieke onderwerpen wordt geschreven. De bedoeling is dat

hierdoor een globaal beeld van het congres als geheel ontstaat, al blijft er natuurlijk, op basis van de interesses van de betrokkenen, sprake van een selectie van datgene wat binnen de diverse divisies aan de orde is geweest. Aan de auteurs is gevraagd vooral in te gaan op algemene trends die binnen hun divisie zichtbaar waren. Van de 11 AERA-divisies komen er 7 aan bod. Zij bestrijken het overgrote deel van hetgeen in het Nederlandse taalgebied onder 'onderwijskunde' of 'onderwijswetenschappen' wordt verstaan.

Literatuur

White, R. (1995). International scholarship and the AERA. *Educational Researcher*, 24(6), 19-21.

Divisie B: Curriculum Studies (M. Mulder, Universiteit Twente)

Ook dit jaar sprong de veelheid aan bijdragen binnen de Divisie Curriculum weer in het oog. In totaal waren er zo'n tachtig sessies bestaande uit verschillende presentaties, alsmede een ledenvergadering en een bestuursvergadering. Er waren 270 voorstellen voor de deadline ingediend. Als de voorstellen die na de deadline bij de organisatiecommissie zijn binnengekomen worden meegeteld, dan ligt dat aantal zelfs boven de 300. Uitsluitend de voorstellen die bijtijds zijn ontvangen zijn beoordeeld. Daarvan waren er 167 ingediend als traditionele paperpresentaties, er waren 41 voorstellen voor symposia en 62 voor experimentele sessies. Het beleid van de AERA is erop gericht het indienen van voorstellen voor de experimentele, meer interactieve sessies te bevorderen. Gezien het aantal van dat soort ingediende voorstellen kan worden opgemerkt dat dit beleid in de praktijk ook redelijk wordt gerealiseerd. Interessant is te vermelden dat de acceptatieratio dit jaar lag op 58%. Dat komt overeen met 157 voorstellen van voldoende, goede en uitstekende kwaliteit. Er waren echter slechts 52 slots in het programma, maar door te onderhandelen met andere Divisies en Special Interest Groups, konden toch alle goedgekeurde voorstellen worden opgenomen in het programma.

Wat de inhoud van het programma betreft, kan worden opgemerkt dat er vier subthema's

waren:

- curriculumonderzoek in de context van de classesituatie;
- curriculumtheorievorming;
- curriculumontwerp, -implementatie en -evaluatie;
- curriculumgeschiedenis, -bewegingen en -praktijken.

Het is ondoenlijk om een overzicht te geven van alle bijeenkomsten binnen de divisie, maar enkele titels van sessies zullen duidelijk maken dat de koppeling tussen de vier genoemde subthema's en de inhoud van de sessies niet altijd duidelijk is.

- Het doordenken van de levens van vrouwelijke leerkrachten: poststructuralistische en materialistische feministische benaderingen.
- Reflecties over de formuleringen van Dewey over democratie en onderwijs: spel als een bron van gemeenschap voor kinderen.
- Vrouwen en curriculum: een feministische exploratie.
- Het herstructureren van onderwijsonderzoek in een problematische democratie.
- Een nieuwe generatie onderzoekers: exploraties in lesbische, homo- en bisexuele problemen in het onderwijs.
- Voor wie en hoe is de vakinhoud van belang: Afrikaans-Amerikaanse leerlingen en het schoolcurriculum.

Deze voorbeelden geven tegelijkertijd aan dat er binnen de Divisie Curriculum grote belangstelling is voor actuele sociaal-maatschappelijke en wetenschapstheoretische problemen en voor de doordenking van de consequenties van die problemen voor de inhoud van het onderwijs voor kinderen en jongeren.

Binnen de divisie is in verschillende andere bijeenkomsten ook aandacht geschonken aan het algemene thema van de conferentie: "Onderzoek voor onderwijs in een democratische samenleving". Zo stond de door de Divisie Curriculum geïnitieerde plenaire bijeenkomst in het teken van verbeelding (imagination), mogelijkheid (possibility) en curriculum, een tamelijk abstracte aanduiding voor een thematiek waarin de rol van de inhoud van het onderwijs voor de vorming van kinderen en jongeren binnen een samenleving, waarvan het democratisch gehalte onder druk staat, centraal

wordt gesteld. Twee eminenties op het terrein van de onderwijskunde, Greene en Sarason presenteerden hun zienswijze op deze problematiek. Beide sprekers gingen in op het belang van onderwijs in de 'arts' (zoals schilderen, muziek, drama en dichtkunst) voor de ontwikkeling van de verbeelding van kinderen en jongeren. Greene betoogde dat deze ontwikkeling van cruciaal belang is. Veel kinderen en jongeren voelen zich nutteloos, wanhopig en zijn het slachtoffer van haat, aldus Greene. Kunstonderwijs kan bijdragen aan de transformatie van hun perceptie van de werkelijkheid en kan een mogelijkheid bieden zich een andere werkelijkheid voor te stellen en aan de realisatie daarvan te werken. Het kunstonderwijs kan aldus kinderen en jongeren een perspectief bieden van waaruit zij betekenis in plaats van betekenisloosheid zoeken, en apathie en geestelijke anesthesie bestrijden. Deze zienswijze gaat terug op de filosofie van Dewey en moet volgens haar niet worden opgevat als het creëren van luchtkastelen: het kunstonderwijs moet direct gerelateerd worden aan de realiteit. Sarason onderstreepte het belang van deze zienswijze en was het op opvallend weinig punten oneens met Greene, ofschoon hij in de herziene versie van zijn 25 jaar oude boek 'Revisiting the Culture of the School and the Problem of Change' weinig woorden wijdt aan het kunstonderwijs en de bijdrage daarvan aan de ontwikkeling van de verbeelding en de mogelijkheid om betekenis te zoeken, bewondering te wekken en de bestaande situatie te transfigureren naar een meer ideale situatie. Sarason wees wel op het belang van de wijze waarop het kunstonderwijs in de praktijk wordt vormgegeven. Hij onderscheidt daarbij productieve en non-productieve onderwijscontexten. De non-productieve vormen voegen weinig toe aan de ontwikkeling van de verbeelding maar komen helaas wel het meest voor in het Amerikaanse onderwijs. Helaas – aldus Sarason – is dit de afgelopen jaren niet verbeterd, vooral ook omdat het onderwijs alle kenmerken mist van zelfcorrigerende en zelfverbeterende systemen die naar kwaliteit streven. En helaas hebben de betrokkenen bij de vormgeving van het onderwijs nog niet het voorstellingsvermogen om het onderwijssysteem te zien als een zelfcorrigerend systeem.

Ofschoon er een kern van waarheid zit in het betoog van Sarason, is de constatering mis-

schien niet geheel juist. Want in een andere, ook zeer druk bezochte bijeenkomst werd gesproken over KEYS, een grootschalig project waarin scholen gebruik kunnen maken van een gestandaardiseerde methodiek voor kwaliteitsbevordering van het onderwijs. Er wordt gebruik gemaakt van een diagnostisch instrumentarium, een kwaliteitscommissie in de school zorgt voor de uitvoering van deze diagnose, de National Educational Association (NEA) verwerkt de gegevens en rapporteert aan de school aan de hand van 39 indicatoren. De presentaties in deze bijeenkomst werden verzorgd door Schneider van NEA, Fullan van de Universiteit van Toronto, Lieberman van Teachers College, Columbia University en Rolles, eveneens van NEA. Professioneel gebruik van een systeem als KEYS kan bijdragen aan de zelfcorrectie en -verbetering waar Sarason reeds jaren voor pleit. De titel van deze bijeenkomst luidde: 'Het gebruik van leerkrachtenonderzoek om lerende organisaties te creëren en de onderwijs-professie te bevorderen'. Dit thema sluit aan bij het algemene beleid van de AERA om onderzoek dat is uitgevoerd door of samen met leraren ook op te nemen in het AERA-programma. Deze trend is overigens al een aantal jaren geleden ingezet. Er is al een aantal jaren plaats voor het presenteren van onderzoek dat is uitgevoerd door of met leraren, maar dit jaar is geprobeerd de leraren nog meer bij de conferentie te betrekken dan in voorgaande jaren. Was dit in het verleden binnen de Divisie Curriculum voornamelijk beperkt tot de rol van de leraar in het implementeren van curriculummaterialen en werd de leraar beschouwd als een lerende, en als consument van onderzoek, en lag de nadruk op implementatie- en 'utility'-onderzoek, nu heeft de AERA-organisatie ook gekozen voor een afzonderlijke dag waarop speciale aandacht is gegeven aan de praktijk. Deze dag heeft een speciale naam meegekregen: 'Sessies met de scholen'.

Ofschoon dit streven met sympathie moet worden bejegend, leidt het wel tot een combinatie van verschillende soorten onderzoek. De reactie van wetenschappelijk onderzoekers is dan al gauw: de kwaliteit van het onderzoek gaat hard achteruit. Er werd bijvoorbeeld in één van de sessies een presentatie gehouden over een nascholingsprogramma op het gebied van archeologie. Docenten van verschillende disci-

plines en onderwijsniveaus konden participeren in het programma. Het programma bestond uit een periode van veldwerk van drie weken in een woestijnachtige omgeving. Het veldwerk bestond onder andere uit het verrichten van opgravingen. Er was een eenvoudig kamp waar de deelnemers konden verblijven. In de presentatie die werd verzorgd door de organisator van deze nascholingsactiviteit, werd uiteengezet welke activiteiten de deelnemers verrichtten en op welke wijze de deelnemers gebruik konden maken van de activiteiten in hun onderwijs. De deelnemers zijn na afloop van het nascholingsprogramma een tijd gevolgd om na te gaan op welke wijze zij gebruik gemaakt hebben van hetgeen zij in het nascholingsprogramma hebben geleerd. Een dergelijke presentatie kan interessant zijn voor docenten en nascholers in dit vakgebied, maar voor wetenschappelijk onderzoekers is het niet veel meer dan een praktijkverhaal zonder wetenschappelijk kader, en kan dan ook moeilijk worden opgevat als een presentatie over wetenschappelijk onderzoek. Misschien moeten de bezoekers van toekomstige AERA's zich meer en meer realiseren dat 'educational research' binnen AERA-verband een rekbaar begrip is waarin uitdrukkelijk plaats is voor praktisch situationeel bepaald onderzoek dat wordt verricht door leraren. Als de AERA als beleid heeft leraren doelbewust bij de conferentie te betrekken, dan is het begrijpelijk dat dit soort onderzoek ook een plaats krijgt binnen het programma. Het is niet ondenkbaar dat leraren daar het meest aan hebben. Maar er ontstaat dan wel een mix van wetenschappelijk onderzoek voor onderzoekers en ideosyncratisch onderzoek voor nascholingsdoeleinden. De toekomst zal uitwijzen of deze combinatie vruchtbaar is. In ieder geval is het bezoekersaantal van de AERA voldoende groot om iedereen een plaats te bieden.

Een interessant fenomeen in dit geheel is dat er dit jaar voor het eerst ook sessies binnen scholen konden worden bezocht. Voor deze bijeenkomsten moest men zich van tevoren opgeven en een eindje reizen, maar het was voor sommigen wel een mogelijkheid een school in New York van binnen te zien.

Zoals bijna ieder jaar was er binnen de Divisie Curriculum weer aandacht voor de vergelijking van de ontwikkelingen in de curriculumtheorie en het curriculumonderzoek in de

Verenigde Staten en Europa. De titel van de eveneens druk bezochte sessie over dit onderwerp was veelbelovend: 'Curriculumtheorie en -onderzoek revisited na 20 jaar. Ligt Europa nog steeds aan kop?' In 1976 is er een bijeenkomst binnen de Divisie Curriculum geweest over hetzelfde onderwerp, waarin is geconcludeerd dat Europa verder was in de curriculumtheorievorming en het curriculumonderzoek dan de Verenigde Staten. De presentaties in deze bijeenkomst werden verzorgd door Hamilton, Weiner, Sander en Calgren (met Kallós). Discussianten waren Kliebard en Yates.

Vanuit de Britse bijdragen werd duidelijk dat inmiddels anders wordt gedacht over talrijke onderwijsproblemen. Veel van de openbare discussies zijn gericht op meer technische aspecten in het onderwijsbestel en op het bevorderen van de effectiviteit van scholen. Volgens de sprekers is dat een beperkte discussie en moet er veel meer aandacht zijn voor de maatschappelijke problemen en de relatie daarvan met het curriculum.

De bijdrage uit Duitsland ging over het falen van het twintig jaar geleden geïntroduceerde nieuwe paradigma voor het onderwijs. Dit paradigma bestond uit een nieuwe visie op de Duitse didactiek, waarbij nadruk kwam te liggen op open communicatie, open curricula en open curriculumontwikkeling. In tegenstelling tot wat werd aangekondigd als een renaissance van de Duitse didactiek moet volgens Sander nu worden geconcludeerd dat de Duitse didactiek dood is en dat zelfs wederopstanding onmogelijk is. Als voornaamste oorzaak voor dit falen wordt aangevoerd dat leerlingen uiteindelijk toch geen macht hebben over hun situatie in het onderwijs en velen zich daarom ook afvragen waarom zij zich druk zouden maken over de democratische discussie. Ook de unificatie en de economische recessie heeft geen goed gedaan aan de verspreiding van de ideeën over de Duitse didactiek. In het Duitse onderwijsbestel (zowel in west als oost) is veel verzet tegen ideologisch getint onderwijs.

Carlgrén gaf een schets van de disconnectie van het onderwijsonderzoek van het onderwijsbeleid en de onderwijspolitiek, respectievelijk de 'National Board of Education' en het Ministerie van Onderwijs in Zweden. De positie van het onderwijsonderzoek is ernstig ver-

slechterd nu het geen integraal deel meer uitmaakt van het onderwijsbeleid en de onderwijspolitiek. Het onderwijsbestel is meer gedecentraliseerd en het onderzoek is momenteel op zoek naar nieuwe doelgroepen. Opvallend is dat de decentralisatie heeft geleid tot centrale besturing van het onderwijs op decentraal niveau en dat er sprake is van een tegenstelling tussen meer instrumentele en meer kritische benaderingen op decentraal niveau. De uitdaging voor onderzoekers is volgens Carlgren om het overleg op decentraal niveau tussen de verschillende categorieën betrokkenen beter te structureren.

Nadat de discussianten vanuit hun perspectieven wat grote lijnen hadden getrokken kwam de prangende vraag in de titel van deze sessie uiteindelijk uit de zaal. Wat was nu het antwoord: ligt Europa nog steeds voor? Na enkele ogenblikken doodse stilte sprak Kallós het verlossende woord: 'Ja'. Helaas gaf hij vervolgens geen argumenten voor zijn antwoord. Gelukkig werd in de zaal wat genuanceerder gedacht over dit antwoord. Het antwoord moet natuurlijk zijn: 'dat hangt er van af'. Het hangt af van het onderwerp binnen de curriculumtheorie en het curriculumonderzoek en uiteraard ook van het perspectief van waaruit de vraag wordt beantwoord. Bovendien waren slechts drie landen uit Europa vertegenwoordigd. En waren er geen presentaties waarin de Europese en Amerikaanse ontwikkelingen systematisch werden vergeleken. Dat zou mijns inziens het antwoord op deze vraag moeten zijn. Als men dat wil weten, moet men dat stelselmatig gaan vergelijken en zijn oordeel niet baseren op onafhankelijke analyses van de ontwikkelingen in drie landen.

Divisie C: Learning and Instruction

C - I: Leren en instructie (L. Verschaffel en E. De Corte, Katholieke Universiteit Leuven)

In Divisie C werd heel veel aandacht besteed aan de thema's die ook gedurende de voorgaande jaren het onderzoek rond probleemoplossen en leren sterk in de belangstelling stonden, m.n. de rol van metacognitie en zelfregulatie in (leren) denken en probleemoplossen, de interactie tussen cognitieve en motivationele aspecten, de waarde en haalbaarheid van

de constructivistische visie op leren en onderwijs, de waarde van het 'situated cognition' paradigma en de diverse interpretaties ervan, het ontwerpen van krachtige leeromgevingen. Onze globale indruk is dat er ten aanzien van deze thema's geen noemenswaardige nieuwe ontwikkelingen te melden vallen. Wel vond er een aantal boeiende confrontaties en discussies plaats en werden tal van empirische gegevens gerapporteerd die de reflectie over en het onderzoek van deze onderwerpen verder vooruit helpen. Omdat we zelf vooral sessies gevolgd hebben waarin vragen omtrent leren en instructie toegespitst werden op het vakgebied 'wiskunde', zullen we ons verslag ook daarop concentreren. Doch vooraf gaan we - exemplarisch - in op enkele sessies waarin de bovengenoemde topics vanuit een meer algemeen perspectief werden behandeld.

In het druk bijgewoond symposium 'Conceptions of learning and understanding', waarin o.a. ook bijdragen van bekende Europese onderwijspsychologen als Entwistle, Säljö en Marton werden gepresenteerd, hield Perkins een opmerkelijke lezing waarin hij zich sterk afzette tegen de klassieke 'representational view of understanding', die uitgaat van de assumptie dat 'understanding' gebaseerd is op de constructie en het gebruik van kennis in de vorm van schematische of propositionele netwerken, regelsystemen e.d. Vertrekend vanuit de idee dat het hebben van begrip van of inzicht in iets enkel maar kan blijken uit 'flexible performance', trachtte Perkins de toehoorders ervan te overtuigen dat het beschikken over en het gebruik maken van (bewuste) representaties noch een noodzakelijke noch een voldoende voorwaarde is om te kunnen spreken van 'understanding'. Volgens hem dient de 'representational view of understanding' vervangen te worden door een 'performance view of understanding'; datgene wat geconstrueerd wordt wanneer er kennis en begrip van iets wordt opgebouwd moet opgevat worden in termen van 'flexible performance capacities'. In hetzelfde symposium lichtte Entwistle zijn ideeën toe over 'knowledge objects' als wezenlijke bestanddelen van de menselijke cognitie en werd door Marton gewezen op de epistemologische moeilijkheden die veroorzaakt worden door het individu van zijn omgeving los te koppelen; deze loskoppeling genereert een

kloof die logisch niet meer overbrugbaar is. In zijn discussiebijdrage reageerde Salomon (fel) op Perkins' betoog; gebruikmakend van voorbeelden zowel uit het dagelijks leven als uit de recente onderzoeksliteratuur, pleitte hij voor het behoud van het representatiebegrip dat door Perkins zo scherp onder vuur was genomen. Zo argumenteerde hij dat het gelijkstellen van 'iets begrijpen' met 'begrijpend gedrag vertonen' en het in vraag stellen van de rol van mentale representaties bij het begrijpen van de werkelijkheid, ons wel zeer (gevaarlijk) dicht brengt bij een behavioristisch georiënteerd gedachtengoed.

In zijn inleiding tot het symposium met de veelbelovende titel 'Situated cognition versus cognitivism: a grounded debate' stelde diSessa de vraag wat de voor- en nadelen zijn van het ijveren voor een paradigma-wisseling binnen een wetenschapsgebied en van het beklemtonen van verschillen tussen de concurrerende paradigma's. Sherin stelde een empirisch onderzoek voor rond (leren) oplossen van fysica-problemen dat als prototypisch voor het traditionele cognitivistische paradigma kon gelden en pleitte op grond daarvan voor het behoud van dit paradigma. Daarna was het de beurt aan Greeno. Eerst zette hij de - inmiddels bekende - verschillen tussen het 'cognitivistische' en het 'situationele' perspectief nogmaals in algemene termen op een rijtje. Daarna deed hij een - niet erg geslaagde - poging om te illustreren hoe deze verschillen (duidelijk) tot uiting kwamen in de wijze waarop (kwalitatieve) onderzoeksgegevens in beide onderzoeksparadigma's geanalyseerd en geïnterpreteerd worden. Tenslotte ging diSessa zelf diepgaand in op de problemen die zich stellen bij de aanvang van een individueel interview. Hoewel zijn bevindingen en interpretaties op zich erg interessant waren, was het voor de (meeste) toehoorders niet zo duidelijk waarin de bijdrage van deze lezing tot het 'grounded debate' over het cognitivistische en het situationele paradigma precies gelegen was.

Mede door het voorkomen van meer radiCALE en meer gematigde standpunten en meer cognitieve en meer sociale varianten van het constructivisme, blijft de constructivistische visie op leren nog steeds het voorwerp van veel discussie. Dit kwam in talloze sessies tot uiting. Nogal wat sprekers verwezen in hun bijdrage

naar het recente maar inmiddels bekende artikel van Phillips van Stanford University 'The good, the bad, and the ugly: The many faces of constructivism' (Phillips, 1995). Het 'ugly aspect', zegt Phillips, is dat het constructivisme in vele gevallen verworpen is tot een soort religie, tot een ideologie. Er is ook een slechte kant ('bad') aan, namelijk het feit dat er een tendens in zit naar absoluut relativisme. Tenslotte is er evenwel een goede kant ('good'), m.n. de klemtoon op de noodzaak van actieve participatie van de lerende en de erkenning van het sociaal karakter van leren.

In een zeer groot aantal sessies stond de rol van metacognitieve kennis en van zelfregulatie in het (leren) probleemoplossen en het (leren) leren centraal. Daarbij ging opvallend veel aandacht naar de mate waarin de metacognitieve denkbeelden en het zelfregulerend vermogen van subjecten (leerlingen, leerkrachten...) te beïnvloeden is. Dit was o.m. het geval in een paneldiscussie getiteld 'Self-regulation in education: future directions in research and applications', waaraan tal van vooraanstaande onderzoekers in dit gebied deelnamen zoals Schunk, Corno, Pintrich, Pressley, Winne en Zimmerman (Weinstein was afwezig). Naast de vraag hoe de verworven kennis gebruikt kan worden voor het verbeteren van het onderwijzen en leren van zelfregulatievaardigheden, kwamen in deze paneldiscussie ook de volgende thema's aan bod: welke zijn de 'component processes' van zelfregulatie, welke zijn de prioritaire vragen en problemen voor het onderzoek in de nabije toekomst, en welke vernieuwingen en verbeteringen in de onderzoeksmethodologie zijn er nodig? Interessant was ook het door Tobias georganiseerde symposium over 'Research on metacognitive judgements and assessments', waarin ingezoomd werd op de laatste vraag, nl. de wetenschappelijke waarde en de praktische bruikbaarheid van eenvoudige instrumenten voor het meten van 'knowledge monitoring' zowel op macro- als microniveau.

In een symposium over 'The design of learning environments: International perspectives' (georganiseerd als een 'joint session' door Division C en het AERA International Relations Committee) werden door Goldman (namens de Cognition and Technology Group at Vanderbilt), Mandl en Vosniadou (Hatano was

afwezig) leeromgevingen voorgesteld die gebaseerd zijn op recente inzichten over effectief leren en onderwijzen. Tijdens de discussie werd door Collins gewezen op de grote diversiteit in de principes die aan het ontwerp van deze verschillende omgevingen ten grondslag liggen. Derhalve werd ervoor gepleit om in de nabije toekomst de discussie en het onderzoek toe te spitsen op het vergelijken van de diverse systemen en het zoeken naar de écht kritische dimensies van krachtige leeromgevingen.

Evenals de voorbije jaren was ook op dit AERA-congres het aanbod op het gebied van het leren en onderwijzen van wiskundige begrippen, vaardigheden en houdingen zeer ruim vertegenwoordigd. Zelfs wie uitsluitend in dit thema geïnteresseerd was, diende bijna voortdurend een keuze te maken uit twee of zelfs drie parallel geprogrammeerde sessies. Met uitzondering van de spectaculair toegenomen belangstelling voor de problematiek van 'authentic assessment' - waarover meer in de bijdrage van Knoors en Dochy - vallen weinig nieuwe ontwikkelingen te melden in vergelijking met bijv. de stand-van-zaken geschetst in onze bijdrage tot de AERA-kroniek van 1994.

Dit betekent dat de onderzoekers van het reken/wiskunde-onderwijs sedert enkele jaren massaal overgestapt zijn op construerend onderzoek, gericht op het ontwikkelen en uitproberen - in nauw contact met practici - van nieuwe vormen van reken/wiskunde-onderwijs. Zo zijn onderzoekers zoals Carpenter, Fuson, Silver, Resnick, Cobb en Lampert allemaal nauw betrokken bij de optimalisering van het nationale reken/wiskunde-onderwijs via het construeren, implementeren en evalueren van nieuwe curricula, deelleergangen, leermaterialen, interventietechnieken, evaluatie-instrumenten, opleidings- en navormingsprogramma's. In al deze studies worden de inmiddels zeer bekende *Standards* (National Council of Teachers of Mathematics, 1989) - een document dat qua strekking en impact vergelijkbaar is met de Nederlandse *Proeve van een nationaal programma voor het reken/wiskunde-onderwijs op de basisschool* (Treffers, De Moor & Feijs, 1989) - uitdrukkelijk en expliciet als referentiekader gebruikt. Bij het overzien van dit onderzoek hebben wij de indruk dat daar noch vanuit theoretisch oogpunt noch

vanuit praktisch oogpunt heel erg veel nieuws te rapen valt voor wie bekend is met de rijke Europese traditie inzake onderzoek en onderwijs van de wiskunde. Een bijkomend probleem is dat al deze Amerikaanse vernieuwingsprogramma's in afzonderlijke sessies voorgesteld en besproken worden, zonder dat ingegaan wordt op eventuele interessante accentverschillen wat betreft achterliggende visie en/of concrete uitwerking.

De socio-culturele context van wiskunde-leren en -onderwijzen blijft een zeer actueel onderzoeksterrein. Ten eerste werden er weerom heel wat papers en posters gepresenteerd waarin ingegaan werd op het informele rekenen van kinderen en volwassenen in buitenschoolse contexten ('ethnomathematics') en op de relatie daarvan met en de betekenis ervan voor het schoolse rekenen. Tijdens deze sessies werd echter nog maar eens duidelijk hoe moeilijk het is om de resultaten van deze op zich boeiende analyses van eigen-aardige manieren van (leren) rekenen te benutten bij de optimalisering van het geïnstitutionaleerd wiskunde-onderwijs.

Internationaal vergelijkend onderzoek is een tweede categorie van onderzoek waarin het socio-culturele aspect van wiskunde-(onderwijs) tot uitdrukking komt. Naast de 'klassieke' comparatieve studies van de wiskundige oplossingsvaardigheden en leer- en denkprocessen van Amerikaanse en Aziatische leerlingen en studenten, werd ook vergelijkend onderzoek gepresenteerd waarin andere landen en culturen betrokken waren. Een goed voorbeeld daarvan was het symposium georganiseerd door Reys over 'Assessing the development of number sense: Collaborative initiatives in four countries', waaraan naast Amerikaanse en Chinese ook Australische en Europese onderzoekers deelnamen. Doch veel verder dan een gezamenlijk referentiekader voor bestudering van 'number sense' en een bruikbare itempool was men nog niet gekomen.

De socio-culturele context van het rekenonderwijs kwam tenslotte ook ruimschoots aan bod in een aantal symposia, papers en posters rond 'the culture of the mathematics classroom'. In dit verband verwijzen we naar ons eigen symposium rond 'Real world knowledge and mathematical modeling of word problems

in schools', met bijdragen van De Corte, Greer, Reusser, Säljö en Verschaffel en een discussiebijdrage van Kilpatrick over de spanning tussen het oplossen van vraagstukken in de schoolse context enerzijds en het oplossen van reële problemen anderzijds.

Tenslotte kwam ook de rol van informatietechnologie (computers, interactieve video) aan bod in heel wat presentaties en demonstraties rond wiskundeleren en -onderwijzen. Bijzonder boeiend was een sessie georganiseerd door Staub, waarin een tiental computerondersteunde 'representational tools for mathematics instruction' (met o.m. bijdragen van Reusser, Derry, Resnick, Greeno, Nathan) werden gepresenteerd en gedemonstreerd. In zijn discussie wist Kaput de verschillende systemen te plaatsen op een aantal kerndimensies (mate van controle, mate van ingebouwde intelligentie, flexibiliteit van de verbanden tussen de diverse representaties). Tevens formuleerde hij een aantal belangrijke uitdagingen waarvoor ontwikkelaars van dergelijke computerleeromgevingen staan. Als belangrijkste wees hij erop dat men cruciale beslissingen dient te nemen over welke aspecten van het wiskundig handelen men kan automatiseren en welke aspecten men de leerling of student best nog zelf laat/leert doen.

Literatuur

- Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, 24(7), 5-12.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Treffers, A., Moor, E. de, & Feys, E. (1989). *Proeve van een nationaal programma voor het reken-wiskundeonderwijs op de basisschool. Deel 1. Overzicht einddoelen*. Tilburg: Zwijzen.

C - II: Onderwijstechnologie en ontwerpkunde (J. Elen en J. Lowyck, Katholieke Universiteit Leuven)

Alhoewel onderwijstechnologie en ontwerp-kunde duidelijk aan elkaar zijn gerelateerd, wordt in de verslaggeving eerst ingegaan op de onderwijstechnologie en vervolgens op de ont-

werpkunde. Afsluitend worden enkele trends genoemd die vanuit de, weliswaar beperkte, selectie uit het aanbod onderkend kunnen worden.

Onderzoek binnen de *onderwijstechnologie* lijkt thans sterk georiënteerd op telecommunicatie. Het begrip 'telecommunicatie' is echter een algemene en omvattende term waaronder zowel het raadplegen van 'homepages' als het uitwisselen van boodschappen over 'het net' vallen. Zo beschouwd verwijst telecommunicatie naar een stadium in de technologische evolutie waarin vrijwel alle gangbare onderwijsleerfuncties via genetwerkte computers (zullen) kunnen worden gerealiseerd. Niet alleen netwerken, maar ook de beschikbare software voor multimedia kwam frequent aan de orde. Veel van de gepresenteerde papers benadrukken, zij het op een zeer descriptieve en hypothetische wijze, de invloed van deze ruime informatie-, spel- en leerbronnen op de positie van de school als instituut. Vooral de relatie tussen 'school and community' en 'school and home' zijn voltreffers in het prospecteren van 'the school of the future'. Een voorbeeld hiervan was te vinden in de paper van Mann die wees op de enorme vlucht die de multimedia industrie neemt inzake 'edutainment'. Op de AERA-conferentie was nog duidelijker dan voorheen de kloof merkbaar tussen enerzijds goed gefocuseerd hypothesetoetsend onderzoek en anderzijds ruwe prospecties over de toekomst van het onderwijs en de rol van de onderwijstechnologie daarin. Blijkbaar krijgt de laatste trend op de AERA-conferentie langzamerhand de meeste aandacht.

Laten we de meer speculatieve bijdragen buiten beschouwing, dan lijkt het onderwijs-technologisch onderzoek zich op een dubbel spoor te bevinden.

Enerzijds bestudeert men, in de lijn van vroeger onderzoek, het belang van symbolen-systemen (tekst, beelden, beweging) waarmee de werkelijkheid wordt gerepresenteerd. De klemtoon ligt dan op het concretiseren en manipuleren van informatie met het oog op het verwerven van kennis (Entwistle spreekt, in dit verband over 'knowledge objects'). Parallel hiermee wordt ervoor gepleit om ook in onderwijsleersituaties het (leren) manipuleren van '(kennis)objecten' centraal te stellen. Enyedy

bijvoorbeeld rapporteerde onderzoek naar de effecten van een computeromgeving op het domein van de waarschijnlijkheidsleer. Hij constateerde dat manipuleerbare representaties bijdragen tot de bespreekbaarheid ervan en dat studenten door representaties geholpen worden om actief met die abstracte materie om te gaan. Schnotz op zijn beurt heeft systematisch onderzoek gepresenteerd over bewegende beelden. Hij stelde vast dat het beste leerresultaat wordt bereikt door leerlingen bij wie visueel-spatieële relaties in figuren overeenkomst vertonen met semantische relaties tussen gepresenteerde kenniselementen. Dit helpt verklaren waarom dynamische figuren een beter leereffect sorteren indien ook de te verwerven kennis procesmatig van aard is.

Anderzijds kwam nieuw onderzoek aan bod over implicaties van de nieuwe technologische mogelijkheden voor het leren. Interessant in dit verband is het onderzoek dat zowel aan de University of Illinois als aan het Education Development Center is verricht naar mentale modellen die leraren construeren over 'Internet' en naar de metaforen die ze hierbij hanteren. Toch blijven de bevindingen op dit vlak eerder opsommend en beschrijvend, zonder specifieke aandacht voor het ontstaan of de evolutie van die modellen of voor de invloed ervan op het gebruik van netwerken in onderwijsleersituaties. In de lijn van de trend naar 'netwerken', passeerde een aantal communicatie-omgevingen de revue die leerlingen/studenten in staat moeten stellen met behulp van CMC (computer mediated communication) op afstand met elkaar samen te werken. Zo stelt de communicatie-omgeving KIE (Knowledge Integration Environment), ontwikkeld in Berkeley en gepresenteerd door Bell en Davis, leerlingen in staat verschillende soorten van evidentie op te zoeken, te rapporteren en vervolgens te bespreken. Doordat leerlingen in KIE ook expliciet worden aangespoord hun opmerkingen en theoretische uitgangspunten te bespreken, wordt verwacht dat ze op een dieper niveau met de materie bezig zijn en hierdoor een dieper inzicht verwerven. Kaye ging nader in op de kenmerken van effectieve netwerk-omgevingen en meer bepaald op 'Model Nets' als systeem voor computer netwerking vanaf de kleuterschool tot het einde van het secundair onderwijs (K-12). Helaas werd geen serieuze evaluatie

van de leereffecten gerapporteerd, maar wel de optimistische conclusie dat schooldistricten die een dicht netwerk creëren, ook bijdragen tot de verbetering van de kwaliteit van hun onderwijs. Het gehanteerde criterium voor deze voorbarige conclusie was uitsluitend de mate van penetratie van computer netwerken. Dat het gebruik van telecommunicatie wel de organisatie van het onderwijs beïnvloedt maar nauwelijks effecten te zien geeft op individuele leerresultaten werd onder andere geïllustreerd in een paper van Hawkins en Honey 'Social factors for networking'. Op grond van ervaringsgegevens werd gewezen op het belang van passende ondersteuning. Omdat deze ondersteuning door computers gemedieerd wordt, dient volgens de onderzoekers deze typische wijze van communiceren vanaf de start in rekening te worden gebracht. Deze gemedieerde ondersteuning riep verder nog vragen op over de behoefte van persoonlijke aanwezigheid van de ondersteunende instantie en over de mate waarin gemedieerde ondersteuning als authentiek en betrouwbaar wordt ervaren door lerenden. Een hamvraag in het startend onderzoek naar effecten van telecommunicatie, is de evaluatie. Zowel Stecher als Quellmalz en Kozma stelden de vraag naar de wijze waarop deze nieuwe mogelijkheden kunnen worden geëvalueerd. Laatstgenoemden presenteerden een complex schema dat echter niet wezenlijk afwijkt van een gangbaar evaluatieschema. Dit wijst op de moeilijkheid om nieuwe omgevingen grondig te analyseren in hun typische maar ladende kenmerken. Volgens Stecher wordt degelijke evaluatie ook bemoeilijkt door de complexiteit die ontstaat door een 'mix' van activiteiten en technologieën, die dankzij de telematica parallel kunnen verlopen op verschillende locaties die wel aan elkaar gerelateerd zijn maar niettemin onafhankelijk van elkaar kunnen opereren.

Een volgende stap in de implementatie van nieuwe technologieën is het aanwenden van 'virtual reality' (VR), waarover een symposium werd georganiseerd. Het systematisch onderzoek naar de mogelijkheden en grenzen ervan voor het onderwijs staat duidelijk nog in de startblokken. Wel zijn tal van evaluatieonderzoeken gaande met soms positieve uitkomsten. Dede ging in op het gebruik van VR voor het bevorderen van onderwijs in de na-

tuurwetenschappen. Hiertoe werden diverse virtuele werelden gecreëerd waarin lerenden zich zelfstandig in een driedimensionele virtuele simulatie-omgeving kunnen bewegen om verschillende hypothesen uit te testen. Het gebruik van VR wordt door Dede beargumenteed vanuit de mogelijkheden tot driedimensionele representatie, het gebruik van meerdere sensorische systemen, en tot 'immersion' en motivatie, eigen aan VR. Terecht stelde hij echter dat onderwijskundig nut hier voorrang verdient, wat niet steeds evident is. VR wordt dan niet zozeer als een omvattend antwoord, maar als een hoogtechnologische en specifieke mogelijkheid naar voren gebracht. De kosten voor het ontwikkelen van VR-omgevingen bracht wellicht mee dat in dit symposium verder alleen de NASA en de Landmacht van de VS aan het woord kwamen. Loftin (NASA) ging nader in op een project waar astronauten werden voorbereid op hun missie voor het herstellen van de Hubble-telescoop. Daarbinnen werd ook tussen astronauten samengewerkt die zich op ver van elkaar gelegen plaatsen bevonden. Goldberg beklemtoonde dat het gebruik van VR voor het leger kostenbesparend kan zijn door de behoefte aan training te reduceren en te verhinderen dat reële uitrusting moet worden gebruikt. Voor hen vormt VR een uitbreiding van gangbare simulaties. In de laatste twee papers werd gewezen op zogenaamde 'simulator sickness' en werd de vraag gesteld naar de mate waarin VR de werkelijkheid getrouw moet afspiegelen om voor de lerende betrouwbaar over te komen.

Met betrekking tot *ontwerpkunde* vallen onmiddellijk twee zaken op. Ten eerste is 'instructional design' als apart onderzoeks- of discussieterrein vrijwel verdwenen. De vroegere debatten over de relatie constructivisme-objectivisme zijn nagenoeg uitgedoofd. Voor zover we konden nagaan was er slechts één groep van paperpresentaties waar in de titel expliciet naar ontwerpkunde werd verwezen en waar overigens weinig nieuws te rapen viel. Zo rapporteerde Maxwell over eigen ervaringen naar aanleiding van zijn poging om binnen een MBA-programma ontdekkingsleren te implementeren. Wedman rapporteerde een studie die het belang van 'analogical problem solving' reveleerde via hardop denken. Vooral de gehan-

teerde methode voor het analyseren van de hardop-denken protocollen lijkt veelbelovend. Er moet worden geconcludeerd dat het onderzoek naar ontwerpen als proces stilgevallen is en het onderzoek naar meer formele ontwerpproducten ondergeschikt is gemaakt aan domeinspecifieke of media-gerelateerde vraagstellingen. Toch dienen nog twee symposia te worden vermeld die dit beeld enigszins relativeren. Vanuit de 'SIG Structural Learning' werd een Amerikaans-Europees symposium georganiseerd over het automatiseren van het ontwerpen. Dit is zowat een traditie geworden op de AERA conferentie, zoals uit de programma's van vele jaren blijkt. Er werd voornamelijk aan taakanalyse aandacht besteed. Zo presenteerde Scandura een gecomputeriseerd hulpmiddel om taken te analyseren en rekening houdende met de uitkomsten ervan opleidingen te ontwerpen. Tennyson demonstreerde een recente versie van GAIDA, GUIDE genaamd, en bepleitte het construeren van een set van losse design tools, in tegenstelling tot een meer geïntegreerd systeem zoals de 'ID-Expert' van Merrill c.s., dat het ontwerpen van courseware volledig zou automatiseren. Op de vraag naar de behoefte aan onderlinge samenhang van een set losse ontwerptools werd door Tennyson niet ingegaan. Dijkstra stelde een systeem voor om doelen en leertaken te categoriseren, als een degelijke basis voor het nader analyseren van leertaken. Schott beklemtoonde eveneens het belang van taakanalyse en lichtte een in Dresden ontwikkelde cognitieve taakanalyse-procedure toe.

Directer met de ontwerpkunde verbonden was het symposium van de University of California in Berkeley. Binnen de onderzoeksgroep waaraan ook Lynn en diSessa participeren wordt expliciet gestreefd naar het formuleren van een prescriptieve instructietheorie. Gesteld werd dat de theorie van Bruner vervangen dient te worden door een theorie die de technologische mogelijkheden van computers mee in rekening brengt. Alle inspanningen binnen de onderzoeksgroep zijn georiënteerd op het formuleren van zo'n instructietheorie. Eerder dan losse fragmenten vormen de verschillende onderzoeken binnen de onderzoeksgroep (bijv. naar communicatie-omgevingen (Bell & Davis), samenwerken (Hoadley & Hsi) en simulatie-omgevingen (Enyedy)) bouwstenen

voor de nog te construeren onderwijstheorie waarin de technologie volledig tot zijn recht komt. Centraal staat hierbij, volgens Gifford, de gedachte dat technologie een integraal bestanddeel kan zijn van een onderwijsleeromgeving en dat leren, ondersteunen van leerprocessen en het ter beschikking stellen van informatie, vooral functies zijn en niet zozeer aan personen of objecten gebonden opdrachten. Wat het kennisdomein betreft, ging veel aandacht gaan naar de mogelijkheden van technologie voor het ondersteunen van de verwerving van wetenschappelijke inzichten.

In een aantal onderzoeken ging aandacht uit naar een vraagstelling die steeds meer met ontwerp-kunde in verband wordt gebracht, namelijk de percepties van studenten en de eventuele invloed ervan op het leren. Hiermee verbonden is de kloof tussen percepties van studenten en onderwijsgevend en ontwerpers. Zo benadrukte Reiman in zijn analyse van onderzoek naar voorkennis, de rol van 'task perception'. Reusser presenteerde een poster over de invloed van percepties over groepswerk en Lowyck rapporteerde tijdens het symposium over het automatiseren van ontwerpen een Leuvense studie (Elen & Lowyck) naar de kennis die studenten bezitten over diverse - ook technologische - leeromgevingen. Inzicht in deze mediërende variabelen kan bijdragen tot effectievere leeromgevingen door met het 'gebruikersstandpunt' meer expliciet rekening te houden tijdens het ontwerpen en uitvoeren van instructie.

In tal van symposia kwamen thema's aan de orde die voor het ontwerpen van onderwijsleeromgevingen belangrijke vragen oproepen. In dit verband kan eerst en vooral verwezen worden naar het symposium over 'Conceptions of learning and understanding' waarover Verschaffel en De Corte reeds gerapporteerd hebben.

Door tal van onderzoekers wordt gezocht naar een nieuwe conceptualisering van factoren die het leren beïnvloeden. Vanuit een sociaal constructivistisch standpunt werd door McCarthy het argument aangebracht dat de determinerende factoren, resp. de voor leren interessante variabelen zich niet bevinden in de (kenmerken van) interagerende actoren maar in de (aard van de) interactie zelf. In dit perspectief

wordt de vraag naar de aard van de interacties die een leeromgeving toelaat belangrijker dan vragen naar de mate waarin een individueel lerende door deze omgeving wordt ondersteund.

Iets breder terugblikkend en op grond van symposia, papers en posters die ook buiten de strikte ontwerp-kunde en onderwijstechnologie vallen, kunnen de volgende tendensen worden onderkend die voor de toekomst van (het onderzoek binnen de) onderwijstechnologie en ontwerp-kunde van belang kunnen zijn.

- * Er is verhoudingsgewijze weinig systematisch onderzoek voorhanden dat beoogt modellen en (deel-)theorieën voor het domein van de onderwijstechnologie en ontwerp-kunde te ontwikkelen.
- * Veel onderzoek in het kader van de onderwijstechnologie richt zich op organisatorische kwesties, en veel minder op leerprocessen en -uitkomsten.
- * Op het micro-niveau van het leren, wordt in toenemende mate een rol toegekend aan biologie en neurologie. Farley baseert zijn T-type - typologie relevant in verband met creativiteit - op de bevinding dat één gen verantwoordelijk zou zijn voor interindividuele verschillen qua 'risk-taking'. In dezelfde lijn beklemtoonde Pintrich in zijn discussie over de stand van zaken in het onderzoek naar voorkennis de behoefte aan 'biological metaphors'.
- * Er blijkt een steeds terugkerende vraag te zijn naar stabiele raamwerken en ankerpunten. Deze zoektocht leidt tot een aanpak waarin de gehele context in zijn complexiteit wordt betrokken, dan wel tot veel aandacht voor persoonsgebonden variabelen ('belief systems', 'task perceptions', 'personality') en hun invloed op het leren. Deze dubbele tendens leidt enerzijds tot een niet-geproblematiseerd gebruik van concepten die de omgeving beschrijven, zoals 'context' en 'culture', en anderzijds tot het beklemtonen van de rol van variabelen als 'personality', 'gender' of erfelijkheid.
- * Er wordt nauwelijks replicatie-onderzoek verricht. In plaats van zich in de lijn van een traditie te plaatsen, gaat de voorkeur uit naar onderzoek dat als nieuw, origineel of als een uiting van de zeer recente paradigma-wisseling wordt bestempeld. Door het af-

wijzen van continuïteit in het onderzoek wordt het erg moeilijk om onderzoeksgegevens te accumuleren tot op het niveau van model- en theorievorming. Het gebrek aan stabiliteit en continuïteit lijkt daarom voor de uitbouw van een solide ontwerp-kunde en onderwijstechnologie problematisch.

- * Oorspronkelijk zeer technische psychologische en/of sociologische begrippen of omschrijvingen van stromingen worden geleidelijk geabsorbeerd en op een algemeen niveau getild waardoor ze al te omvattend en betekenisloos worden. Een dergelijke evolutie vond reeds vroeger plaats, bijvoorbeeld voor de concepten 'self-regulation' en 'constructivism' waarop alle mogelijke betekenissen werden geprojecteerd. Thans lijkt dit het geval voor termen als 'representation', 'mental model', en zelfs 'distributed cognition'.

Bij wijze van conclusie kan gesteld worden dat de AERA-conferentie opnieuw gewezen heeft op de spanning tussen het zoeken naar stabiliteit en algemeen geldende regels enerzijds en het erkennen van de complexiteit en eigenheid van ieder onderzoeksobject anderzijds. Niet voor niets spreekt bijv. Pintrich in een symposium over zelfregulatie over 'a cludge' en 'the rich messiness of learning'.

Divisie D: Measurement and Research Methodology (E. Knoors, Katholieke Universiteit Leuven en F. Dochy, Open universiteit Heerlen)

Het zal duidelijk zijn dat Divisie D te veel omvat om alles hier te beschrijven. Wel kunnen we enkele algemene trends aangeven. Daarna geven we een overzicht van enkele door ons bijgewoonde presentaties en workshops.

Algemene trends

In totaal werden er in deze divisie ongeveer 75 sessies georganiseerd. Hierin is de laatste jaren een steeds groter aandeel te zien van onderwerpen in verband met 'assessment' en 'measurement', gepaard gaande met een krimp van de zuiver methodologische en psychometrische sessies.

Veel voorkomende onderwerpen zijn 'Performance Assessment', 'Measurement of

Change', 'Assessment-Based Educational Reform', 'Embedding Assessment systems', 'Collaborative assessment', 'Assessment of Growth', 'Growth modeling', 'New forms of Assessment', etc. Ook het slechts één keer opduiken van de term 'betrouwbaarheid' in de programma-index, tegenover de vele sessies over 'validiteit', vormt een aanduiding voor deze ontwikkelingen. We lichten deze trend kort toe.

De laatste jaren is er sprake van een verschuiving van een test-cultuur naar een 'assessment'-cultuur; van een (puur) psychometrische aanpak naar een aanpak waarin de aandacht vooral uitgaat naar de authenticiteit van het testmateriaal; van een aanpak waarin primair de aandacht uitgaat naar betrouwbaarheid, naar een aanpak waarbij validiteit en 'fairness' centraal staan. Door de al te grote aandacht voor de 'high-stakes testing programs' (dit wil zeggen, prestatietoetsen met belangrijke consequenties voor de leerling) is men tot het inzicht gekomen dat toetsing in eerste instantie voor de student bedoeld is. Hij moet het nut ervan inzien en er iets mee kunnen. Daarnaast heeft onderzoek de grote invloed aangetoond van toetsing op het leren van studenten en hun studiegedrag in het algemeen. Daarmee gepaard ging de stijgende bewustwording van het niet-authentieke karakter van examens en het verlies aan geloof hierin als valide toetsen van leerresultaten. Klassieke testen blijken vaak niet eerlijk te zijn voor de student. De reactie hiertegen wordt steeds sterker. Het is belangrijker valide, authentieke toetsen te ontwikkelen die eerlijk zijn voor studenten en die via meerdere methoden worden afgenomen. Men beklemtoont steeds meer het belang van formatieve toetsen, het bewaken van de voortgang van het leerproces, en van het geven van duidelijke, informatieve feedback.

Workshops

Naast presentaties werd er ook een aantal minicursussen en 'pre-conference courses' over onderzoeksmethoden georganiseerd. Op basis van hun jarenlange ervaring met PLSPath verzorgden Keeves en Campbell een 'preconference session' van één dag getiteld "Soft Modeling - Path Analysis 'Lite'". Hierin werd een software programma voor pad-analyses gepresenteerd - PLSPath - dat zich van andere on-

derscheidt doordat het gebaseerd is op een 'least squares' aanpak en vrij is van onderliggende assumpties. Het programma werd voorgesteld als krachtig maar eenvoudig genoeg om met een minimale training en voorkennis te gebruiken. Ten minste, als je weet hoe je de resultaten moet interpreteren. Een minpunt van het programma is de grote hoeveelheid output en het feit dat het om een DOS-programma gaat. Duidelijk was dat ook de instructeurs geen gebruik maakten van alle detailgegevens die de analyses opleveren, maar zich enkel beperken tot de 'jackknife' procedures.

Een drie uur durende cursus over 'Interrater Reliability in Performance-based Assessment Measures' werd verzorgd door Abedi. Over het reduceren van variantie, veroorzaakt door de scoorders, werd weinig nieuws verteld: Men dient concrete en valide scoringsrubrieken te voorzien en de scoorders moeten goed getraind zijn. Voor het schatten van de interbeoordelaarsbetrouwbaarheid werden acht relatief eenvoudige methoden voorgesteld, variërend van een Product-Moment Correlatie tot een 'William's index of agreement'. De vergelijking van deze diverse waarden voor 'interrater reliability' geeft een grotere mate van zekerheid omtrent de consistentie tussen scoorders, ook al vertrekken deze methoden van een verschillende basis. Abedi legde ook uit dat het gebruik van meerdere methoden om 'interrater reliability' te schatten van belang is, omdat ze beïnvloed worden door verschillende aspecten van de studie, zoals het aantal scoorders en de grootte van de steekproef.

Opvallend, en in contrast met de relatief eenvoudige maten uit Abedi's sessie, is dat in een sessie over 'Performance assessment scoring issues' diverse meer complexe methoden en analyses werden gepresenteerd om verschillen tussen beoordelaars aan het licht te brengen. Twee presentaties in deze sessie verdienden zeker de aandacht. Zowel Martin als Clauser gaven een goede presentatie over het geautomatiseerd scoren van complexe 'performances'. Het is een goede ontwikkeling dat er verder nagedacht wordt over het automatiseren van het scoren van 'performance assessment'. Voor toetsontwikkelaars, en zeker ook voor de opdrachtgevers daartoe, is de hoeveelheid tijd en moeite die gepaard gaat met het scoren van 'performance assessment' veelal een belem-

mering, vooral als men dergelijke toetsvormen op grote schaal wil toepassen. Martin en Clauser maakten dan ook duidelijk dat het automatiseren van het scoren van 'performance assessment' alleen zin heeft als de studentenaantallen groot zijn. In hun geval ging het over een artsenexamen dat voor tienduizenden subjecten moest dienen.

'Performance assessment – Alternative assessment'

Als we de titels van alle sessies in deze divisie en de onderwerpindex van het programmaboek bekijken, zien we dat 'performance assessment' op AERA '96 erg in de belangstelling stond. Bij een tiental sessies was 'performance assessment' in de titel opgenomen, terwijl niet minder dan 38 sessies in de index onder 'performance assessment' gerangschikt stonden. In de presentaties over dit onderwerp kunnen eigenlijk twee hoofdthema's onderscheiden worden. Een eerste thema betreft de relatie tussen toetsing en instructie: In welke mate kunnen toetsen richting geven aan het onderwijs in de klas, hoe kan toetsing geïntegreerd worden met het curriculum, en welke ervaringen zijn er met betrekking tot het invoeren van nieuwe toetsvormen? Een tweede thema betreft de vraag hoe aan dergelijke instrumenten voor 'performance assessment' vorm gegeven moet worden: Welke toetsvorm is geschikt, hoe moeten de items eruit zien, hoe zit het met de betrouwbaarheid en validiteit van 'performance assessment'?

Alvorens enkele interessante presentaties in verband met deze twee thema's te beschrijven, wijzen we er nog op dat het onderzoeks- en ontwikkelingswerk in verband met 'performance assessment' heel sterk toegespitst is op het gebied van wiskunde en taal. In het vervolg van dit stuk zullen we exemplarisch ingaan op 'performance assessment' binnen het gebied wiskunde, omdat de meeste door ons bijgewoonde presentaties daarop betrekking hadden.

De enorme belangstelling voor de kwestie van 'assessment' bij onderzoekers en ontwikkelaars van het wiskunde-onderwijs (zie de bijdrage van Verschaffel & De Corte) houdt verband met het groeiend besef dat elke vernieuwing van dit onderwijs tot mislukking is gedoemd, indien die niet gepaard gaat met of zelfs geleid wordt door parallelle ontwikkelin-

gen inzake toetsontwikkeling en -gebruik. In verscheidene sessies had men het in dit verband over 'assessment-based educational reform'.

Een van de opvallendste karakteristieken van 'performance assessment' is dat items vaak ingebed zijn in authentieke of realistische situaties. In een sessie getiteld 'Investigating relationships between student learning and assessment in primary schools', stelde Cooper zich echter de vraag of dit de eerlijkheid van de toetsing ten goede komt. Op basis van een pilootstudie met een aantal 10- en 11-jarige leerlingen beargumenteerde hij dat de sociale klasse waartoe kinderen behoren, verschillen teweeg kan brengen in de manier waarop zij toetsitems interpreteren en begrijpen. Op basis van de resultaten van interviews die hij met de leerlingen hield enerzijds en gebruikmakend van theorieën van Bernstein en Bourdieu anderzijds, kwam hij tot de (voorlopige) conclusie dat kinderen uit de arbeidersklasse hun oplossingsstrategieën en antwoorden meer baseren op contextgebonden ervaringskennis dan kinderen uit de middenklasse, die gebruik maken van abstracte, wiskundige kennis. Het interessante aan Coopers studie is dat daaruit blijkt dat het inschakelen van deze concrete ervaringskennis deze leerlingen hindert bij het oplossen van wiskundige toepassingsproblemen, omdat zij die aanwenden op plaatsen waar eigenlijk wiskundige kennis vereist is. In die zin vormt Coopers analyse een interessante aanvulling op de studies die in het symposium over 'Real-world knowledge and mathematical modelling of word problems in schools' (zie de bijdrage van Verschaffel & De Corte) werden besproken. Cooper droeg aan het slot van zijn betoog dan ook de waarschuwing uit dat wanneer deze conclusies in verder onderzoek bevestigd worden, toetsontwikkelaars voor moeilijke beslissingen komen te staan omtrent het gebruik van toetsopgaven die aan de ene kant het onderwijs kunnen verbeteren, maar aan de andere kant leiden tot minder eerlijke toetsresultaten.

In dezelfde sessie wees ook Harlen op mogelijke negatieve invloeden van 'assessment' op het leren. Om juist een positieve invloed van 'assessment' op het leren te bewerkstelligen, is het volgens Harlen belangrijk dat het formatief toetsen zo gedefinieerd wordt, dat het duidelijk te onderscheiden is van summatief toetsen. Pas

als de essentiële verschillen blootgelegd zijn, kunnen voorstellen gedaan worden om beide toetsvormen aan elkaar te koppelen. De assumpties (a) dat formatieve toetsen *niet* wezenlijk verschillen van summatieve toetsen, en (b) dat de resultaten van een formatieve toets ook *zonder meer* geschikt zijn voor een summatieve evaluatie, zijn namelijk bronnen voor een negatieve impact van toetsen op het leren. Men moet zich realiseren dat de informatie die leerkrachten opdoen tijdens hun onderwijsactiviteiten fragmentarisch en vaak tegenstrijdig is. Deze incongruente gegevens zijn voor een formatief doel goed om aan te duiden wat een leerling vooruit helpt of juist remt in zijn leerproces. Harlen benadrukt echter dat voor een summatieve evaluatie weliswaar hetzelfde materiaal gebruikt kan worden, maar dat de gegevens in het licht gesteld moeten worden van de criteria die voor alle studenten gelden. Dit betekent dat het complete werk van een leerling holistisch gezien moet worden, waarna beoordeeld wordt of het daaruit voortvloeiende profiel voldoet aan de criteria.

Met betrekking tot het tweede thema, namelijk de vormgeving van toetsinstrumenten, werd veelvuldig de vraag gesteld welk type opgaven geschikt is om leerprocessen en vaardigheden van hogere orde in beeld te brengen. In het algemeen was er een trend waar te nemen tegen gestandaardiseerde toetsen en, meer algemeen, tegen geschreven, korte-antwoord items. In het symposium 'Developing valid large-scale mathematics performance assessments that provide diagnostic information' beschreef Close - net als Cooper in de eerder genoemde sessie - hoe in Engeland de toetsontwikkelaars zich afkeerden van dergelijke toetsen en de weg insloegen naar meer 'authentieke' toetsen met complexe testitems; doch onder druk van de regering dienden ze terug te keren naar de klassieke 'paper-and-pencil' toetsen. Van den Heuvel-Panhuizen verzette zich dan weer tegen auteurs als Joffe en Clements, die toetsen met geschreven, korte-antwoord opgaven per se overboord willen gooien. Zij beargumenteerde dat dergelijke toetsen niet uitgebannen kunnen en moeten worden. Men zou volgens haar beter energie kunnen steken in het verbeteren van deze toetsvorm, zodat de negatieve kenmerken en invloeden ervan geëlimineerd worden. Daartoe moe-

ten wel een aantal taboes doorbroken worden die lange tijd de vormgeving van dergelijke toetsen bepaald hebben, zodat men bijvoorbeeld aanvaardt dat een wiskundig probleem meer dan één goed antwoord kan hebben en dat het antwoord op een probleem soms niet exact te bepalen is. (In dit verband merkte Lesh in zijn discussiebijdrage op dat "there is no right way to be good in math"). Dan blijkt dat ook geschreven, korte-antwoord opgaven en zelfs meerkeuzevragen zeer informatief kunnen zijn over de leerprocessen, strategieën en hogere denkvaardigheden van leerlingen.

Ook in de sessie getiteld 'Issues in math assessment' stond het tweede thema, namelijk de vormgeving van 'performance assessment', centraal. Zoals eerder opgemerkt, is de tijd die gemoeid is met 'performance assessment' veelal een probleem. Misschien is het daarom dat Suzuki zich de vraag stelde wat het minimum aantal items is in 'performance based tests'. Met behulp van diverse statistische technieken is ze de betrouwbaarheid van toetsen met verschillende lengten nagegaan, en op basis daarvan concludeerde zij dat reeds met negen opgaven een betrouwbare toets samengesteld kan worden; mits ze zorgvuldig gekozen worden kunnen vijf items zelfs al voldoende zijn. Hoewel het op zich goed nieuws is dat ook een toets met zo weinig opgaven al een betrouwbare meting kan opleveren, slaagde Suzuki er onvoldoende in te expliciteren wat de aard van dergelijke opgaven moet zijn. Bovendien merkte Ankenmann als discussant op dat een toets met zo'n klein aantal items altijd in te boeten heeft aan validiteit. Eerder is reeds aangeduid dat de validiteit van toetsen in een stijgende belangstelling staat, tegenover een overwaardering van de betrouwbaarheid in het verleden. Deze eenzijdige aandacht voor de betrouwbaarheid van een toetsinstrument, waarbij de inhoudsvaliditeit verwaarloosd wordt, is ook Freudenthal een doorn in het oog geweest. Door Brown werd in dezelfde sessie een studie gepresenteerd die zowel de betrouwbaarheid als de validiteit van het door hem gebruikte instrument voor 'performance assessment' van wiskunde, in ogenschouw neemt. Een nog maar eens bevestigde, maar nog te weinig toegepaste bevinding van deze studie is, dat analyses gebaseerd op de item-responstheorie en de klassieke testtheorie heel goed gecombi-

neerd kunnen worden in de toetsconstructie. Tevens kan deze combinatie van methoden aangewend worden om het toetssysteem in het oog te houden en te verbeteren, onder andere met betrekking tot de betrouwbaarheid en de validiteit. Ons inziens wordt het vaak voorgesteld alsof de item-responstheorie en de klassieke testtheorie in een soort machtsstrijd verward zijn. Wellicht is het blijvend probleem dat een praktijktoepassing van de IRT goed gecalibreerde items vereist. Theoretici propageren de ene of de andere psychometrische methode, terwijl in de studie van Brown de twee een symbiose aangaan. FACETS (het gebruikte IRT model) schaaft alle facetten van 'performance assessment' (studenten, scoorders, items, enzovoort), waardoor het mogelijk is om elementen binnen één facet te vergelijken, alsook verschillende facetten onderling te vergelijken. Door de combinatie van FACETS en de klassieke testtheorie kan de overeenstemming tussen scoorders, de interbeoordelaarsbetrouwbaarheid en de toetsvaliditeit worden onderzocht.

Divisie G: Social Context of Education

(G. W. Meijnen, Universiteit van Amsterdam)

Traditioneel vormen onderwerpen als 'ouderschool', de relatie tussen schoolloopbanen en sociaal-economische respectievelijk etnische herkomst en de relatie met sekse het hart van de divisie. Afgeleide onderwerpen als effecten van tweetalig onderwijs, verwachtingen van leerkrachten, leerlingen en ouders, processen van schoolkeuze en sociale relaties binnen klassen behoren uiteraard ook tot de veelvuldig opgevoerde onderwerpen.

Opvallend is voorts dat Amerikaanse onderzoekers in toenemende mate belangstelling aan de dag gaan leggen voor wat zich in het onderwijs in Japan, China en andere snel expanderende economieën afspeelt. Daarbij gaat het niet alleen om het niveau van de leeropbrengsten maar ook om bijvoorbeeld seksespecifieke curriculuminhouden en om de relaties ouderschool. Uit de aandacht voor de laatste onderwerpen blijkt dat men ook op zoek is naar de inbedding van de instructiecontext in deze landen. Daardoor zal waarschijnlijk duidelijk worden wat wel en niet te gebruiken is voor de Amerikaanse context.

Traditioneel wordt binnen deze divisie ook tijd ingeruimd voor papers die meer beschouwend en al of niet maatschappijkritisch van aard zijn. De kwaliteit van deze lezingen was ook dit jaar zeer wisselend, variërend van erudiete analyses tot pamfletachtige teksten. Kennelijk lukt het de organisatoren niet om zodanige selectieprincipes te hanteren dat een minimumniveau gegarandeerd is.

Hoogtepunten dit jaar voor divisie G waren een optreden van Bernstein en een bijeenkomst gewijd aan de onlangs overleden Coleman. In een tot de nok toe gevulde zaal presenteerde Bernstein op de hem eigen wijze, te weten het prototype van de hoogleraar uit 1950, een nieuwe theorie omtrent elites, onderwijsconcepties en organisatieprincipes. Het leverde een groot aantal omschrijvingen van categorieën op met een zekere herkenningswaarde, maar daar bleef het dan ook bij. De wetenschappelijke verantwoording van de taxonomie was nogal aan de magere kant. Wellicht zal hij de toekomst besteden aan de uitwerking ervan. De op het podium genodigde referenten waaronder Apple en Wexler bereden of hun eigen stokpaardjes of bleven erg vaag in hun respons.

De bijeenkomst gewijd aan Coleman was eveneens zeer druk bezocht. Een aantal prominente collega's waaronder Bidwell en Sörensen refereerden de betekenis van Coleman voor zowel het onderwijssociologisch onderzoek als voor het beleid. Beide zijn onomstreden. Toch ontbraken enige kritische noten niet: zo presenteerde Sörensen een heranalyse van door Coleman gebruikte data om de relevantie van sociaal kapitaal aan te tonen en kwam daarbij tot geheel andere conclusies.

Hoewel qua organisatie niet maar qua onderwerp wel tot Divisie G behorend, was het symposium naar aanleiding van het boek van Herrnstein c.s. *The Bell Curve*, de moeite waard. De drie referenten waren allen van mening dat de inhoud van het hoofdstuk over de beleidsaanbevelingen niet logisch uit de analyses voortvloeit. Een der referenten noemde de wetenschappelijke analyse dan ook het sausje dat de ideologische boodschap verteerbbaar moest maken. Wellicht lag het aan de keuze van de panelleden dat het verwachte verbale vuurwerk uit de zaal achterwege bleef.

Samenvattend: het blijft binnen deze divisie een moeilijke opgave bijeenkomsten te selecteren die de moeite waard zijn als het niet om zeer gevestigde namen gaat. En ook dan kan men teleurgesteld worden. Een oplossing ligt niet voor het oprapen.

Divisie H: School Evaluation and Program Development (B. P. M. Creemers, Rijksuniversiteit Groningen)

De eindredacteuren van de kroniek van de AERA-conferentie 1996 verzochten me een bijdrage te leveren over de sessies in Divisie H. Divisie H van de AERA moet het bewijs vormen dat wetenschap en onderzoek ook dienstbaar kunnen zijn voor praktijk en beleid. Het is dan ook een divisie waarin praktijkgericht onderzoek – dat betekent de presentaties van onderzoekers die in de onderwijspraktijk werkzaam of van practici die ook wat onderzoek doen – aan de orde komt. Een dergelijke divisie is bij het plannen van mijn programma voor de AERA-conferentie niet mijn primaire vertrekpunt, maar omdat ook onderwijseffectiviteit en de verbetering van onderwijs gebaseerd op de resultaten van het onderzoek waarnaar mijn belangstelling uitgaat in deze divisie hun plaats krijgen, was het doornemen van het programma op onderdelen van Divisie H zinvol.

Wanneer men het programma van Divisie H en dat van de 'Special Interest Group School Effectiveness and School Improvement' in beschouwing neemt, dan lijkt het erop dat onderwijs-effectiviteit weer terug is op de agenda van de AERA. Er waren meer presentaties over deze onderwerpen dan in voorafgaande jaren, toen het erop leek dat onderwijs-effectiviteit na een hausse in de belangstelling in de jaren tachtig, in de jaren negentig nauwelijks meer enige interesse kon opwekken. Maar opname onder het kopje 'School Effectiveness and School Improvement' garandeert nog niets over de inhoud noch over de kwaliteit van het gebodene. Bij veel presentaties en in de programmering, zowel bij de SIG 'School Effectiveness and School Improvement' als bij Divisie H, ligt het accent op de onderwijsverbetering, hoewel bij de SIG nog een relatie met effectiviteitsonderzoek aanwezig is.

Over de relatie tussen onderwijs-effectiviteitsonderzoek en onderwijsverbetering is de

afgelopen jaren door diverse auteurs gepubliceerd. Dit gebeurde omdat er, na de aanvankelijk innige relatie, tot uitdrukking komend in deze SIG en in de International Congress for School Effectiveness and Improvement, toch enige distantie tussen beide groepen is ontstaan. Teleurstelling was er aan de kant van degenen die betrokken waren bij onderwijsverbetering vanwege het feit dat er geen pasklare recepten en antwoorden voor problemen in de onderwijspraktijk te geven waren en ook omdat het gewoon af en toe niet lukte; bij degenen die betrokken waren bij onderzoek op dit terrein was er een grotere interesse, naar mijn mening terecht, in de verdere theoretische fundering en beproeving van een integraal model voor onderwijseffectiviteit dan in het opleveren van deze pasklare antwoorden voor de problemen in de onderwijspraktijk, die bovendien vaak ook niet gebruikt werden.

Op basis van de analyse van verschillende perspectieven van effectiviteitsonderzoek en onderwijsverbetering zijn door diverse auteurs voorstellen ontwikkeld voor verbetering van het gebruik van elkaars activiteiten (Creemers & Reezigt, in druk; Reynolds, Hopkins & Stoll, 1993).

In de presentaties op de AERA werd de nadruk gelegd op schoolverbetering, waarbij veelvuldig klachten werden geuit over het onderwijseffectiviteitsonderzoek. Vooral het feit dat het onderzoek te ver van de praktijk afstaat en deeloplossingen aanbood die niet geïntegreerd konden worden in de onderwijspraktijk, de benadrukking van de kwantitatieve methoden voor onderzoek en het belang van de opbrengsten van onderwijs in termen van leerprestaties, vormden daarbij punten van kritiek. Onderwijsverbetering lijkt vooral hand in hand te gaan met andere dan cognitieve opbrengsten en een accentuering van de processen in het onderwijs. Het succes van de onderwijsverbetering wordt dan afgemeten aan de mate waarin veranderingen zijn opgetreden in de factoren en karakteristieken van onderwijseffectiviteit, zoals schoolleiderschap, teamvorming en dergelijke, in plaats van aan de leerwinst van leerlingen.

De presentaties over de onderwijsverbetering betroffen vaak aspecten en factoren die eventueel in de verbetering een rol zouden kunnen spelen, zoals het verlangen van leer-

krachten om te participeren in besluitvormingsprocessen gedurende het veranderings- ('restructuring')-proces (Taylor, Tashakkori & Hardwick, 1996). Bij deze studie waren slechts zeven scholen betrokken, waarbij niet een relatie tussen de wens tot participatie en de participatie zelf kon worden geconstateerd, maar 'gelukkig wel' een duidelijke relatie tussen participatie en positieve opvattingen over participatie. Ook in andere onderzoeken van onderwijsverbetering valt het geringe aantal eenheden op. Er werden verstrekkende conclusies getrokken over zaken als school-based management en onderwijs in grotere tijdseenheden op grond van een gering aantal waarnemingen. Waarschijnlijk wordt dit voor een deel veroorzaakt door de praktijkgerichtheid - er moet een praktijkadvies verstrekt worden - en de betrekkelijk beperkte werksituatie van de onderzoeker.

In de sessies waarin gerapporteerd werd over onderwijseffectiviteitsonderzoek viel in de eerste plaats op dat veel bijdragen van buiten de Verenigde Staten kwamen. Het Nederlandse onderzoek op dit terrein was goed vertegenwoordigd, evenals het Britse onderzoek. Dit zou erop kunnen wijzen dat onderwijseffectiviteit als researchprogramma vooral in Europa tot ontwikkeling gekomen is, terwijl in de Verenigde Staten de nadruk is komen te liggen op onderwijsverbetering, behalve in een aantal groepen die vanouds vooral met effectiviteitsonderzoek bezig zijn (Louisiana State University, Vanderbilt University).

In een door de International Congress for School Effectiveness and Improvement georganiseerde sessie werden diverse onderzoeken van het London Institute of Education gepresenteerd (Goldstein, Myers, Sammons, Stoll, Thomas, Mortimore en West). In dit instituut bestaat een stabiele traditie op het terrein van effectiviteitsonderzoek. De nadruk ligt er op kwantitatief onderzoek, zowel van min of meer integrale modellen als de aspecten ervan, zoals stabiliteit, contexteffecten en het leren en onderwijzen. Er is echter ook aandacht voor meer op onderwijsverbetering gericht onderzoek, zoals dat van Stoll. Het accent ligt daarbij op de evaluatie van onderwijsverbeteringsprojecten en de feitelijke begeleiding van onderwijsverbetering, maar er wordt ook onderzoek verricht naar bijvoorbeeld de kenmerken van ineffec-

tieve scholen en de wijze waarop ineffectieve scholen kunnen veranderen. In het onderzoek van Myers, Stoll en Reynolds (1996) bleken ineffectieve scholen zich te kenmerken door gebrek aan visie, ongericht leiderschap, disfunctionele onderlinge relaties binnen de staf en ineffectieve praktijken in de klas. Het is echter de vraag of dit niet een vorm van achteraf interpreteren is, in plaats van het via empirisch onderzoek vaststellen van verschillen tussen effectieve en ineffectieve scholen. Bovendien lijkt het erop dat, zoals wel vaker gebeurt, ineffectiviteit wordt gedefinieerd in procestermen in plaats van gerelateerd aan de opbrengsten van onderwijs. Het is een vorm van 'outlier' onderzoek, waarbij scholen met elkaar vergeleken worden op kenmerken waaraan 'effectiviteit' wordt toegeschreven.

Onderwijskundig leiderschap is nog steeds een belangrijk onderwerp in Amerikaans onderzoek op het terrein van onderwijseffectiviteit en er werd uiteenlopend onderzoek op dit terrein gepresenteerd, van de ontwikkeling van een schaal voor 'leader authenticity' en de relatie tussen authenticiteit en effectiviteit - in dit geval de gepercipieerde effectiviteit, die dan ook hoog was (Henderson & Brookhart, 1996) - tot de presentatie van meta-analyses zoals van Hallinger en Heck (1996) over de rol van het schoolhoofd in onderzoek uit de periode 1980-1995. In hun ijver om het belang van onderwijskundig leiderschap empirisch te ondersteunen zijn ze zover gegaan dat een heranalyse van het oorspronkelijke materiaal van Van de Grift (Van de Grift, 1990) werd uitgevoerd om duidelijk te maken dat ook Van de Grift wel degelijk effecten van onderwijskundig leiderschap had kunnen vinden wanneer hij zijn materiaal op een andere wijze had geanalyseerd.

De presentaties over de resultaten van onderzoek van de integrale modellen voor onderwijseffectiviteit, waarbij diverse niveaus en factoren, variabelen, op de diverse niveaus in beschouwing worden genomen, waren veelal van Europese oorsprong. Het lijkt erop dat daar de interessante ontwikkelingen, zowel wat deelaspecten betreft, zoals stabiliteitsonderzoek, effectgroottes, afzonderlijke componenten, zoals kwaliteitszorg, evaluatie, instructie-effectiviteit, als de evaluatie van onderzoek van integrale modellen vandaan moeten komen en bovendien - voorlopig - uit een vrij beperkt

gebied: Nederland en Engeland. Ander internationaal onderzoek waarover wel in tijdschriften wordt gepubliceerd ontbrak op de AERA.

Het belang van incorporatie van instructie-effectiviteit in integrale effectiviteitsmodellen bleek uit de resultaten van onderzoek van Rowan, Chiang en Miller (1996). De effectgroottes waren weliswaar niet indrukwekkend in deze studie, maar ondersteunen het belang van de effectiviteit van het onderwijzen.

Een interessante sessie gerelateerd aan onderwijseffectiviteit was die over de gelegenheid tot leren. Dit oorspronkelijk binnen de International Association for the Evaluation of Educational Achievement (IEA) geïntroduceerde begrip om de overlap tussen het curriculum en de toets aan te duiden is langzamerhand aardig in betekenis opgerekt. In de Verenigde Staten is er bovendien bijgekomen dat het een belangrijke politieke lading heeft gekregen in verband met een discussie over het kerncurriculum in relatie tot gelijke kansen. 'Opportunity to learn' ging naast toets-curriculumoverlap ook betekenen: 'content coverage', 'content exposure', 'content emphasis' en in de nieuwe TIMSS-studie (Third International Mathematics and Science Study) is het ook de 'quality of instructional delivery'. Daarmee is 'opportunity to learn' het overkoepelende begrip voor instructie-effectiviteit geworden (Schmidt, Valverde, Houang, Wiley, Wolfe & Bianchi, 1996). Behalve tijd ('time on task') lijkt nu ook de kwaliteit van instructie binnen het begrip 'opportunity' gebracht te worden, hetgeen naar mijn mening een te ver gaande oprekking van het begrip is.

Een interessant onderwijsverbeteringsproject werd gepresenteerd door de Memphis 21st Century Schools Exposition, een groep die ook de organisatie van het volgende congres van de International Congress for School Effectiveness and Improvement in 1997 in handen heeft. In Memphis werd de mogelijkheid geboden aan scholen om in diverse vormen van schoolverbetering te participeren. Scholen konden zelf kiezen op welk programma zij intekenden. Er is een uitgebreid scala van programma's voorhanden, van een soort 'outward bound'-achtig programma tot dat van Slavin: 'Roots and Wings', dat sterk gestructureerd en onderwijsgericht is. Tijdens een bijeenkomst werd informatie over alle programma's gegeven en

scholen konden vervolgens intekenen. Op de AERA werden de resultaten van de keuze en de motieven ervoor gepresenteerd; later zal ook vergelijkend onderzoek naar de effecten plaatsvinden. Het probleem dat zich bij deze vergelijkende evaluatie voordoet is dat scholen geen vergelijkbare achtergrond hebben, zowel wat de doelstellingen betreft als wat de leerlingpopulatie en de effecten van het onderwijs betreft. Uit de eerste resultaten van het onderzoek blijkt dat scholen een programma hebben gekozen dat grotendeels aansluit bij de vigerende praktijk in de school voorafgaande aan het project.

Hoewel het gebodene misschien niet altijd meevalt, zoals uit deze korte beschrijving moge blijken, blijft de AERA-conferentie natuurlijk toch de perfecte gelegenheid om een indruk te krijgen hoe onderzoek en praktijkontwikkeling er op bepaalde terreinen bijstaan. In de weinig gestructureerde warwinkel van opvattingen over schoolverbetering min of meer in relatie met onderwijseffectiviteit, was de presentatie van Miles op de 'business meeting' van SIG 'School Effectiveness and School Improvement' een lichtpunt. Miles is zeker geen groot voorstander van effectiviteitsonderzoek, zeker niet in kwantitatieve zin - hij laat zich ook weinig aan kwantitatief onderzoek gelegen liggen - maar binnen het kwalitatieve onderzoek, waarin aanbevelingen gedaan worden voor de verbetering van de onderwijseffectiviteit, gaat hij logisch, analytisch en argumenterend te werk. Hoewel ik er niet aan moet denken dat zijn droom over de stand van zaken in onderwijseffectiviteit in de 21ste eeuw namelijk: alleen maar kwalitatief onderzoek over onderwijs in verandering en een 'theorie' waarin dit wordt vastgelegd, ooit uit zal komen.

Literatuur

- Creemers, B. P. M., & Reezigt, G. J. (in druk). School level conditions affecting the effectiveness of instruction. Geaccepteerd voor publicatie in *School Effectiveness and School Improvement*.
- Grift, W. van de (1990). Educational leadership and academic achievement in elementary education. *School Effectiveness and School Improvement*, 1(1), 26-41.
- Hallinger, P., & Heck, R. H. (1996). *The principal's role in school effectiveness: An assessment of substan-*

- tive findings, 1980-1995*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.
- Henderson, J. E., & Brookhart S. M. (1996). *Leader authenticity: Key to organizational climate, health and perceived leader effectiveness*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.
- Myers, K., Stoll, L., & Reynolds, D. (1996). *Understanding ineffectiveness*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.
- Reynolds, D., Hopkins, D., & Stoll, L. (1993). Linking school effectiveness research: Items for a research agenda. *School Effectiveness and School Improvement*, 4(1), 37-58.
- Rowan, P., Chiang, F.-S., & Miller, R. J. (1996). *Using research on employee performance to study teaching effectiveness: Results from an analysis of teacher effects on student achievement in mathematics using NELS:88 Data*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.
- Schmidt, W. H., Valverde, G. A., Houang, R. T., Wiley, D. E., Wolfe, R. G., & Bianchi, L. J. (1996). *Studying the intended, implemented and attained curriculum: Strategies for measurement and analysis in large-scale international research*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.
- Taylor, D. L., Tashakkori, A., & Hardwick, J. (1996). *A typology based on teachers' desire for participation in decision making*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.

Divisie J: Postsecondary Education

(J. F. M. J. van Hout, Universiteit van Amsterdam)

Ik ga meestal om de paar jaar naar de AERA-conferentie om te bezien of er nieuwe trends zijn waar te nemen in het Amerikaanse onderzoek van het hoger onderwijs. Dit jaar was ik er bovendien om als discussiant op te treden bij de VOR-sessie die door de divisie Hoger Onderwijs werd georganiseerd. Vanuit beide perspectieven geef ik een bijdrage aan deze kroniek.

Ook dit jaar was het onderwerp 'student drop-out, student attrition, student succes' weer een 'evergreen'. De invloed van Tinto, Pascarella en Terenzini is nog steeds groot, mede door een niet aflatende stroom van eigen onderzoek over dit onderwerp. De beleidsrelevantie van dit onderwerp is hoog. In de USA is er een grote uitval van studenten na het eerste studiejaar: bijna 50% bij 2-jarige colleges, ongeveer 25% bij 4-jarige instellingen en 10% bij de zeer selectieve instellingen (Tinto, 1993). Zeker als men bedenkt, dat in de USA bijzonder veel aandacht wordt besteed aan plaatsing en toelating van studenten, zowel van de kant van de instellingen als van de kant van de studenten en hun ouders, is dat een opmerkelijk feit.

Het meest interessante paper over dit onderwerp vond ik dat van Braxton, Shaw Sullivan en Johnson (Vanderbilt University en Belmont University, Nashville). Zij presenteerden een diepgaand review van onderzoeken die het model van Tinto als theoretisch kader hadden gebruikt. Zij testten daarbij de 15 relaties waaruit het model bestaat en relateerden de resultaten aan de theoretische kritiek op het model. Slechts een beperkt deel van de theoretische relaties in het model wordt door het empirisch onderzoek ondersteund. Hierbij blijken tevens verschillen te bestaan tussen typen instellingen en tussen vrouwelijke en mannelijke studenten. De auteurs verwerpen het model als geheel niet, maar zijn wel van mening dat ingrijpende aanpassing van het model noodzakelijk is, waarbij vooral gebruik gemaakt moet worden van andere theoretische invalshoeken, zoals organisatiekunde, economie en psychologie. Zij zijn tevens van mening dat het model voor 'institutional research' een goed uitgangspunt is.

Een tweede onderwerp dat vaak in Divisie J aan de orde komt is het universitaire personeel, 'faculty' of 'administration'. Onderwerpen zijn dan bijv. onderzoeksproductiviteit, satisfactie, onderwijsoriëntatie. Opmerkelijk was dat een aantal sessies hierover georganiseerd was door 'the Committee on the role and status of women in educational R&D'. De beperkte opname van vrouwen en raciale minderheidsgroepen in de wetenschappelijke staf is in de USA een groot probleem. De succesvolle Amerikaanse onderzoeker is blank en van het mannelijk geslacht. Dit type onderzoek leek

mij ook voor de Nederlandse universiteiten interessant, omdat Nederland ongeveer de laatste plaats in de westerse wereld inneemt wat het aantal vrouwen in de hogere wetenschappelijke rangen betreft.

Naast enkele 'N=1 studies' was het paper van Bonous-Hammarth (UCLA) een goede uitschieter. Uit haar onderzoek blijkt, dat mannen meer publiceren dan vrouwen en blanken meer dan Afro-Americans. Naast enkele persoonskenmerken (bijv. ambitie), omgevingskenmerken (bijv. takenpakket), blijken zowel ras (als zelfstandige factor) en 'person-environment fit' belangrijke predictoren. De 'P-E fit' is een maat voor overeenkomst in waarden en opvattingen van de respondent met (onderzoekers in) de werkomgeving. Zij is ontleend aan onderzoek naar arbeidssatisfactie in andere organisaties en blijkt dus ook relevant voor universiteiten. Vraag is dan wel hoe een 'P-E fit' ontstaat. Dit onderzoek maakte voor mij nog weer eens duidelijk, dat universiteiten niet alleen maar talent en doorzettingsvermogen belonen, maar ook andere criteria en regels hanteren die het voor bepaalde groepen (bijv. vrouwen) moeilijk maken tot de top door te dringen. Bovendien blijkt ook dat problemen die (voorlopig nog) specifiek Amerikaans lijken te zijn, aspecten in zich hebben die in meer algemene zin gelden voor universiteiten en professionele organisaties. De problematiek van raciale minderheden in het wetenschappelijk corps zegt meer over de universiteit zelf, dan ik in eerste instantie had gedacht.

Ieder jaar krijgt de VOR de gelegenheid een sessie op de Annual Meeting van de AERA te vullen met Nederlands onderzoek. Dit jaar was de divisie hoger onderwijs aan de beurt en betrof het Nederlands onderzoek naar studieloopbanen in het hoger onderwijs. Gepresenteerd werd onderzoek van het IOWO (KU Nijmegen), ROA/TO (RL/UT), SCO/KI (UvA) en SEO (UvA). Gezien de aard van de kroniek zal ik deze bijdragen niet inhoudelijk bespreken. Er zijn twee punten die ik wel vermeldenswaard voor deze kroniek vind. De eerste is, dat we ons over de kwaliteit van dit onderzoek geen zorgen hoeven te maken. Zowel op methodologisch als theoretisch vlak deed het niet onder voor onderzoek dat door Amerikaanse collegae over dit onderwerp werd gepresen-

teerd. Dit raakt meteen mijn tweede punt, namelijk het nut van een dergelijke AERA-sessie die door de VOR jaarlijks gevuld mag worden. De papers zijn niet op de gebruikelijke manier geselecteerd, maar zouden bij een anonieme beoordeling door Amerikaanse peers normaal gesproken ook geaccepteerd zijn. Opname in een paper- of postersessie met vergelijkbare Amerikaanse papers had voor de hand gelegen. In dat geval hadden deze papers en presentatoren meer bekendheid bij onze collegae gekregen dan nu het geval was. De VOR-sessie trok namelijk vooral Nederlandse belangstellenden! Zowel voor de kwaliteit van de papers en presentaties, als voor het opbouwen van netwerken lijkt het mij beter als Nederlandse onderzoekers via de koninklijke weg de Annual Meeting van de AERA binnen stappen, dan via een zij-ingang. De papers van divisie J zijn gemiddeld niet van dat bijzondere niveau, dat Nederlands onderzoek daar uit de toon zou vallen.

Literatuur

- Bomous-Hammarth, M. (1996). *Succeeding in the academy: Predicting the influence of value congruence on the research productivity of black and white scientists*. Paper presented at the Annual Meeting of the American Educational Research Association, New York City, New York.
- Braxton, J. M., Shaw Sullivan, A., & Johnson, R. M. (1996). *Appraisal Tinto's theory of college student departure*. Paper presented at Annual Meeting of the American Educational Research Association, New York City, New York.
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Divisie K: Teaching and Teacher Education

(D. Beijgaard en J. Vermunt, ICLON, Rijksuniversiteit Leiden)

Erg veel presentaties binnen Divisie K stonden in het teken van de kennis van leraren en de wijze waarop leraren het best (na)geschoold kunnen worden. Tamelijk veel aandacht werd eveneens besteed aan het beoordelen van leraren (in opleiding). Hieronder wordt vooral op beide eerstgenoemde thema's ingegaan. Enige extra aandacht wordt besteed aan de door de Holmes Group gepresenteerde bevindingen

over de huidige stand van zaken met betrekking tot de opleiding van leraren in de Verenigde Staten.

Het onderzoek naar de kennis van leraren is overwegend kleinschalig en kwalitatief van aard. Het gaat hier vooral om narratief onderzoek of onderzoek waarin het verhaal ('story') van een leraar centraal staat. De argumenten die pleiten voor dit type onderzoek lopen zeer uiteen. Doyle hield bijvoorbeeld een pleidooi voor narratief onderzoek naar aanleiding van door hem gesignaleerde tekortkomingen van het traditionele, empirisch-analytische onderzoeksmodel: 'waarheid' is een communicatieve aangelegenheid en is niet gebonden aan een bepaalde methode. Andere argumenten hebben bijvoorbeeld betrekking op het doorvoeren van veranderingen of vernieuwingen. Elbaz stelde dat het voor het welslagen van veranderingen of vernieuwingen een voorwaarde is om te luisteren naar hetgeen leraren te vertellen hebben over de aan de orde zijnde materie. Zij vindt dat zowel de onderzoeker als de leraar beiden hun voordeel moeten doen met narratief onderzoek. De moeilijkheden van dit soort onderzoek moeten volgens Elbaz niet worden onderschat. Over de functies die een onderzoeker met dit soort onderzoek nastreeft moet duidelijkheid bestaan evenals wiens 'voice' (van de onderzoeker of de onderzochte) uiteindelijk wordt beschreven in het onderzoeksrapport en het methodologische proces op basis waarvan primaire data uiteindelijk worden gerepresenteerd.

Desalniettemin waren er veel presentaties die niet verder gingen dan het beschrijven van unieke situaties met daarin opgenomen veel letterlijke citaten van leraren. Een flink aantal van deze verhaalbeschrijvingen was volstrekt a-theoretisch: zowel een theoretische kader als een terugkoppeling van resultaten naar enige vorm van theorie ontbraken veelal. In een aantal gevallen wordt dit bewust gedaan, bijvoorbeeld vanuit een zeer emancipatoire gedachte (onderzoekers moeten luisteren naar wat de praktijk hun te vertellen heeft) of op basis van een volgestrekte afkeer van theorie. Meestal is echter sprake van onwetendheid, hetgeen als zorgwekkend moet worden aangemerkt. Onderzoekers durven zelfs op basis van een N=1 studie te generaliseren en vinden het vaak niet terecht wanneer ze daarop aangesproken worden.

Inhoudelijk waren de presentaties over de kennis van leraren zeer versnipperd. Interessant was bijvoorbeeld de presentatie van Sherin over het analyseren van de kennis van leraren en het onderscheid tussen 'cognitive modeling' en 'knowledge system analysis'. De eerste verwijst naar het werk van Schoenfeld waarin op basis van een cognitief model fijnmazige en opeenvolgende analyses worden ondernomen van kennis waarover iemand beschikt, de tweede verwijst naar het werk van Shulman en Grossmann die vooral zijn geïnteresseerd in de genese en ontwikkeling van kennis. Ook wees Sherin erop dat het goed is te onderscheiden tussen vorm (structuur en organisatie) en inhoud (waar iets over gaat en waartoe het dient) van kennis. Het meeste onderzoek naar de kennis van leraren is vooral gericht op de vorm ervan.

Verhalen ('stories') van leraren worden vaak gebruikt als een middel voor scholing van leraren (in opleiding). In diverse presentaties werd duidelijk dat zij reflectie bij leraren bevorderen. Andere middelen voor scholingsdoeleinden waarvan een positieve invloed uitgaat op de eigen professionele ontwikkeling zijn bijvoorbeeld conferentie-achtige bijeenkomsten. Tijdens zulke bijeenkomsten, die overigens goed in het totale curriculum moeten worden gepland, worden perspectieven van anderen (peers) ingezet om de complexiteit van de eigen onderwijssituatie beter te begrijpen. Daarnaast werd er tijdens diverse presentaties op gewezen dat (a.s.) leraren ook veel kunnen leren van hun eigen biografie.

In nogal veel presentaties werd de nadruk gelegd op de rol van de praktijk bij de opleiding van leraren. De achtergrondgedachte daarbij is dat ervaring de beste leermeester is. Chin en Russell presenteerden bijvoorbeeld het opleidingsmodel van de universiteit waarbij zij werkzaam zijn. Dit model gaat sterk uit van praktijkleren: na een korte introductie verkeren de studenten lange tijd in de praktijk. De kwaliteit van hun leerproces is daarmee sterk afhankelijk van de begeleiding van leraren in de stagescholen. Op deze en andere modellen waarin een eenzijdig accent op de praktijk ligt, werd flink kritiek geleverd. Bullough, bijvoorbeeld, verwijt zo'n opleiding dan dat die te veel is gebaseerd op een 'craft' model. Leren van ervaring zonder daarbij een perspectief te krijgen

is in zijn ogen onjuist, waarbij onder andere het gevaar bestaat dat leraren in opleiding misconcepties ontwikkelen. Bovendien is zo'n opleiding ethisch onjuist: kinderen kunnen niet zomaar aan 'blanco' leraren worden overgelaten.

In andere presentaties werd te sterk van de praktijk afhankelijke lerarenopleidingen onder meer verweten a-theoretisch te zijn. In dit verband moet worden gewezen op het onderzoek van Nedell. Uit dit onderzoek blijkt bijvoorbeeld dat er een positieve samenhang is tussen professionele ontwikkeling en het beschikken over kaders ('frameworks') om de eigen praktijk te analyseren. Door opleiding verworven kennis kan wel degelijk een belangrijke rol spelen in het proces van leren onderwijzen. Het lijkt erop dat deze opvatting weer enige terreinwinst boekt.

Verandering in de lerarenopleiding werd vooral bepleit naar aanleiding van een door de Holmes Group gehouden onderzoek over ontwikkelingen die zich de laatste tien jaar hebben voltrokken in de lerarenopleidingen in de Verenigde Staten en waar het de komende jaren met deze opleidingen naartoe moet. Er werd vooruit gelopen op het nog in 1996 te publiceren, derde rapport van deze groep, getiteld 'Tomorrow's schools of education'. (Het eerste rapport, 'Tomorrow's teachers', verscheen in 1985; het tweede werd in 1990 uitgebracht onder de titel 'Tomorrow's schools'). Alle lerarenopleidingen die lid zijn van de Holmes Group (N=190) evenals een aantal personen zijn op een aantal voor lerarenopleidingen relevante punten onderzocht of bevraagd, zoals het streven naar goede prestaties, handelingsplannen, onderzoek in de lerarenopleiding, organisatiegraad en leiderschap. Op basis van de onderzoeksresultaten moet worden geconcludeerd dat veranderingen in de lerarenopleidingen nog steeds in de kinderschoenen staan. Gepleit wordt voor een meer omvattend plan van aanpak dan in 1985 werd voorgesteld in het rapport 'Tomorrow's teachers'. In de nabije toekomst moet op basis van dit plan gewerkt worden aan het ontwikkelen van het curriculum voor de lerarenopleiding, het formuleren van voorwaarden om continue professionele ontwikkeling van leraren te realiseren, het ontwikkelen van onderzoeksprogramma's c.q. -agenda's en het houden van debatten over wat een lerarenopleiding nu eigenlijk precies moet zijn.

Diverse bekende wetenschappers was gevraagd te reageren op hetgeen door de onderzoekers van de Holmes Group werd gepresenteerd. Kelley onderkende het algemene effect dat van het werk van de Holmes Group uitgaat. Hij stelde dat niemand is geïnteresseerd in het programma van de lerarenopleiding; het gaat om de verbinding van dit programma met hetgeen in scholen gebeurt. In dit opzicht miste hij enige aandacht bij de Holmes Group voor de evaluatie van leraren via leerresultaten van leerlingen. Lanier ging in haar reactie in op het gegeven dat lerarenopleidingen niet altijd onderzoek doen. Zij acht onderzoek van groot belang voor het ontwikkelen van een kwalitatief goede lerarenopleiding. Hargreaves was in zijn reactie het meest uitgebreid en sprak vooral vanuit de positie van onderzoeker die kritiek levert op een onderzoek. Het doel van een lerarenopleiding is in zijn ogen niet zozeer een goede voorbereiding op het leraarsberoep op zich, maar moet vooral gericht zijn op het verbeteren en veranderen in de scholen zelf. Ook hekelde hij in dit verband de samenwerking van lerarenopleidingen met de zogenoemde 'professionele ontwikkelingsscholen': leraren leren onderwijzen in deze scholen, maar komen daarin na hun afstuderen niet te werken. Het zou de Holmes Group volgens hem niet al-

leen te doen moeten zijn om de status van leraren te verhogen, maar ook om het ontwikkelen van hoge standaarden waaraan leraren zouden moeten voldoen.

Binnen de divisie onderwijzen en lerarenopleiding, tenslotte, was het moeilijk onderzoek te vinden over kennisconstructieprocessen van leraren-in-opleiding. Wel waren er diverse presentaties over verschillende opleidingsdidactische modellen en de veronderstelde effecten ervan. Deze hadden over het algemeen echter weinig oog voor de wijze waarop aanstaande leraren de didactische maatregelen die deel uitmaken van deze modellen interpreteren en gebruiken. Ook de strategieën die zij hanteren om hun kennisbasis op te bouwen en bijvoorbeeld de kloof tussen hun ervaringskennis en theoretische kennis te overbruggen, bleven sterk onderbelicht.

Aan deze kroniek werkten mee: *D. Beijaard, E. De Corte, B. P. M. Creemers, F. Dochy, J. Elen, J. F. M. J. van Hout, E. Knoors, J. Lowyck, G. W. Meijnen, M. Mulder, N. Verloop, J. Vermunt en L. Verschaffel*

De eindredactie werd verzorgd door *N. Verloop en L. Verschaffel*

Boekbesprekingen

R. Damhuis

Interaction and second language acquisition. Participation and control in classroom conversations with young multilingual children

Academisch proefschrift Universiteit van Amsterdam.

IFOTT (Studies in language and language use 19), Amsterdam 1995, 330 pagina's, ISBN 90 74698 19 0

In Nederland is de school voor anderstalige kinderen in kleutergroepen een belangrijke linguïstische omgeving, omdat hun tweede-taalverwerving sterk afhankelijk is van verbale interactie tijdens activiteiten in de klas. De dissertatie van Damhuis bevat het verslag van een in groep 1 en 2 van de basisschool uitgevoerd onderzoek naar de interactie tussen anderstalige kleuters en hun gesprekspartners (Nederlandstalige kinderen en leerkrachten). Uitgaande van de gedachte dat het nemen van initiatieven de tweede-taalverwerving stimuleert, is gekeken naar de *participatie* van het anderstalige kind, dat wil zeggen de initiatieven die het in de conversatie neemt, en de *controle*, dat wil zeggen de relatieve verdeling van de initiatieven over het anderstalige kind en zijn gesprekspartners. Er worden drie soorten participatie en controle onderscheiden: globale participatie en controle betreffen de hoeveelheid spraak van de tweede-taalleerders en hun gesprekspartners, beurtwisselingsparticipatie en -controle hebben betrekking op de rol van de tweede-taalleerders en hun gesprekspartners bij beurtwisselingen, en onderwerpsparticipatie en -controle hebben betrekking op hun rol bij het vaststellen van het onderwerp van de conversatie. De achterliggende gedachte is dat de relatie tussen initiatieven en tweede-taalverwerving een monotoon positieve functie is: hoe meer anderstalige kinderen participeren, en hoe meer ze in controle zijn in de conversatie, hoe meer hun tweede-taalverwerving wordt gestimuleerd. Het probleem met deze aanname is echter dat er noch een sterke theoretische basis, noch veel empirische evidentie voor is.

In hoofdstuk 1 wordt aandacht besteed aan de theoretische achtergronden van het onderzoek, de resultaten van eerder onderzoek op het gebied van participatie en controle, en de factoren die participatie en controle kunnen beïnvloeden. Aan het eind van het hoofdstuk worden de vol-

gende onderzoeksvragen geformuleerd:

1. Hoe zijn de drie soorten participatie en controle verdeeld tussen anderstalige kinderen en hun gesprekspartners in verschillende activiteiten in kleutergroepen?
2. Beïnvloeden de situationele factoren type activiteit, observatiemoment en percentage anderstalige leerlingen in de kleutergroep participatie en controle?
3. Welke praktische implicaties voor de verbetering van gelegenheden tot tweede-taalverwerving in kleutergroepen kunnen worden afgeleid uit de resultaten?

Hoofdstuk 2 bevat een uitgebreide beschrijving van de verschillende aspecten van de onderzoeksoepzet. Bij het onderzoek waren 15 kleutergroepen betrokken van 14 basisscholen. In totaal namen 121 anderstalige en 277 Nederlandstalige kinderen deel. Omdat de scholen een zo goed mogelijk beeld moesten geven van gemiddelde Nederlandse kleutergroepen met anderstalige kinderen, is gestreefd naar een zo groot mogelijke verscheidenheid met betrekking tot de selectiecriteria plaats van de school, leeftijd van de leerlingen, percentage anderstalige leerlingen (dit varieerde van 3.9% tot 81.5%; in absolute aantallen van 1 tot 22) en beschikbaarheid van speciale Nederlands als tweede taal (NT2) lessen (in 8 groepen niet beschikbaar, in 7 wel). Afgezien van het percentage anderstalige leerlingen wordt er weinig aandacht besteed aan de verschillen. Zo is bijvoorbeeld geen gebruik gemaakt van de mogelijkheid om de resultaten in de groepen met en zonder speciale NT2-lessen met elkaar te vergelijken. De interacties zijn opgenomen door een niet-participerende onderzoeker tijdens drie activiteiten: vrije kringgesprekken, werken in een groepje en speciale NT2-lessen. Er zijn op drie momenten opnamen gemaakt: aan het begin, het midden en het eind van het schooljaar. In totaal zijn er 143 opnamen gemaakt en getranscribeerd. De analyse vond plaats op het niveau van de transcripten. Als co-deerseenheid is de interactie-eenheid gekozen, die gedefinieerd is als een eenheid die bijdraagt aan de conversatie. Het begrip eenheid wordt hierbij terecht ruim opgevat: het varieert van hele zinnen tot non-verbale eenheden die van belang zijn voor de interactie. De resultaten van de analyses worden niet per individuele spreker gepresenteerd, maar per sprekertype. Er wordt gerapporteerd over de participatie van een gemid-

deld anderstalig kind en de controle in drie paren van sprekerstypen: alle leerlingen als groep tegenover de leerkracht, een gemiddeld anderstalig kind tegenover de leerkracht en een gemiddeld Nederlandstalig kind tegenover een gemiddeld Nederlandstalig kind. Daarnaast zijn multivariabele regressie-analyses uitgevoerd om de relatie tussen enerzijds participatie en controle en anderzijds de onafhankelijke variabelen – type activiteit, observatiemoment en percentage anderstalige kinderen in de groep – te bepalen.

Hoofdstuk 3 is gewijd aan globale participatie en controle. De hoeveelheid spraak die het anderstalige kind produceerde werd gemeten door het aantal beurten en het aantal interactie-eenheden. De manier waarop ze de spreektijd gebruikten werd gemeten door het aantal interactie-eenheden per beurt en een omvangsindex, die een indicatie geeft van de gemiddelde syntactische complexiteit van de uitingen. De vraag is echter in hoeverre deze omvangsindex een geschikte maat is voor participatie en controle, aangezien hij sterk samenhangt met taalvaardigheid.

De hoeveelheid spreektijd blijkt het grootst te zijn in de speciale NT2-groep, maar de manier waarop deze wordt gebruikt in termen van de beurtlengte en de omvangsindex is het grootst in het werken in een groepje. Alle leerlingen zijn (gedeeltelijk) in controle over de leerkracht in alle drie de activiteiten met betrekking tot de mate van spreektijd, maar de leerkracht heeft steeds controle op het gebied van het gebruik van de spreektijd. Het laatste geldt ook wanneer we de anderstalige leerling met de leerkracht vergelijken. De controle tussen anderstalige en Nederlandstalige leerlingen is gelijk verdeeld, behalve met betrekking tot de omvangsindex, waar het Nederlandstalige kind in controle is. Type activiteit heeft een veel groter effect op globale participatie en controle dan de beide andere factoren (observatiemoment en percentage anderstalige leerlingen in de groep).

In hoofdstuk 4 komen de operationalisatie en de kwalitatieve analyse van beurtwisselingsparticipatie en -controle aan de orde. Hierbij worden de initiatieven tot beurtwisseling die de gesprekspartners nemen bestudeerd aan de hand van een taxonomie die gebaseerd is op werk van Van Lier (1988) met als belangrijkste verschil dat elke uiting zowel voor prospectieve als retrospectieve aspecten wordt gecodeerd. Het

prospectieve aspect geeft aan in welke mate de eenheid van de huidige spreker bepaalt wie de spreker van de volgende unit is. Hoe hoger deze mate, hoe hoger de huidige spreker participeert en hoe meer controle hij heeft over de volgende spreker. Het retrospectieve aspect heeft betrekking op de mate waarin de selectie van de spreker van de huidige eenheid is bepaald door de huidige spreker zelf of door de spreker van de vorige eenheid. Als dit meer in handen is van de huidige spreker, neemt deze meer actief deel en heeft hij meer controle dan de vorige spreker. Damhuis beschrijft op gedetailleerde en heldere wijze de 15 prospectieve en de 27 retrospectieve categorieën van de taxonomie. De prospectieve en retrospectieve aspecten worden op een continuüm geplaatst dat de mate van initiatief van de spreker aangeeft. De uitgebreide voorbeelden geven een aardig beeld van de mate waarin de deelnemers in de verschillende activiteiten initiatief (kunnen) tonen. In het vrije kringgesprek geeft de leerkracht meestal aan wie de volgende spreker mag zijn, maar het kan ook anders, zoals blijkt uit twee groepen waar de kinderen waren getraind om met elkaar te praten en niet via de leerkracht. Ook uit de voorbeelden uit de speciale NT2 groep blijkt dat daar veel afhangt van het gedrag van de leerkracht. Zulke voorbeelden werken voor de op de praktijk gerichte lezer zeer verhelderend.

In hoofdstuk 5 komen de kwantitatieve resultaten met betrekking tot beurtwisselingsparticipatie en -controle aan de orde. Als we participatie bekijken met betrekking tot het *retrospectieve* aspect blijkt het werken in een groepje stimulerend voor tweede-taalverwerking: in deze activiteit toont het anderstalige kind veel initiatieven. Ook in de speciale NT2-groep toont het kind veel initiatieven, naast een grote hoeveelheid responsen. In het vrije kringgesprek vertoont het anderstalige kind zowel weinig initiatieven als weinig responsen. De leerkracht heeft in de meeste gevallen controle, maar weer blijkt het werken in een groepje de meest gunstige controlemogelijkheden voor de kinderen te geven. Het blijkt dat de participatie van het anderstalige kind in *prospectieve* beurtwisseling heel beperkt is. De participatie is wederom het hoogst in het werken in een groepje, lager in de speciale NT2-groep en het laagst in het vrije kringgesprek. Controle wordt in het werken in een groepje gedeeld tussen de leer-

kracht en de leerlingen en tussen de twee groepen leerlingen onderling. In de beide andere activiteiten heeft de leerkracht de controle. De beurtwisselingsindices (die de mate van beurtwisselingsinitiatief van een bepaald sprekerstype aangeven) tonen dat het anderstalige kind de hoogste initiatiefwaarde bereikt in het werken in een groepje en de minste in de speciale NT2-groep. De leerkracht heeft in alle activiteiten op beide indices controle. De multiniveau analyse van de relatie tussen de drie situationele factoren en de beurtwisselingsparticipatie toonde dat alleen type activiteit effect heeft. De multiniveau analyse voor beurtwisselingscontrole geeft aan dat type activiteit effect heeft, maar er zijn ook interactie-effecten (het belangrijkste is type activiteit en percentage anderstalige leerlingen) en er resteert ongemodelleerde groepsvariatie.

Hoofdstuk 6 gaat in op onderwerpsparticipatie en -controle, dat wil zeggen de mate waarin een bepaald type spreker vrij is om het onderwerp van de conversatie te bepalen. Hiertoe is een klassificatiesysteem ontwikkeld dat is gebaseerd op een model van Keenan en Schieffelin (1976) en dat 9 categorieën bevat. Er zijn twee hoofdgroepen: een continue eenheid bouwt voort op het voorafgaande, een discontinue eenheid introduceert een onderwerp dat niet aansluit bij het direct voorafgaande. Ook voor de onderwerpsdimensie is een continuüm opgesteld waarop de categorieën zijn geplaatst naar hun de mate van initiatief: zwakker voor de continue categorieën, sterker voor de discontinue. De kwalitatieve beschrijving van de onderwerpsdimensie in de drie activiteiten geeft het volgende beeld: in het vrije kringgesprek en de speciale NT2-lessen is de mate waarin het kind actief kan deelnemen sterk afhankelijk van de leerkracht: wanneer deze het bepalen van het onderwerp aan de leerlingen overlaat en het kind de tijd geeft om te spreken lijkt de produktie gunstig te worden beïnvloed. In het werken in een groepje wordt het onderwerp vaak, maar niet altijd, bepaald door de uit te voeren taak.

Hoofdstuk 7 bevat het verslag van de kwantitatieve analyse van onderwerpsparticipatie en -controle. Ook met betrekking tot deze dimensie blijken de anderstalige kinderen de meeste initiatieven te ontplooiën bij het werken in een groepje. In de speciale NT2-groep introduceren de anderstalige kinderen weliswaar meer nieuwe onderwerpen dan in de andere activiteiten, maar

omdat ze ook veel minder-initiatiefrijke uitingen produceren is de onderwerpsindex in deze activiteit het laagst. De onderwerpsindex geeft overigens aan dat de gemiddelde initiatiefwaarde voor de anderstalige kinderen in alle activiteiten erg laag is. Het meest opvallende resultaat met betrekking tot onderwerpscontrole is dat niet alleen de leerkracht controle uitoefent op de anderstalige kinderen, maar dat de Nederlandstalige kinderen op deze dimensie tevens controle hebben over de anderstalige kinderen. Een Nederlandstalig kind toont aanmerkelijk meer onderwerpsinitiatief dan een anderstalig kind. Uit de multiniveau regressie-analyse blijkt dat onderwerpsparticipatie en -controle in alle drie de paren van sprekerstypen worden beïnvloed door type activiteit. Hiernaast speelt het percentage anderstalige leerlingen in sommige gevallen ook een rol: een hoger percentage vergroot de controle van de leerkracht.

De hoofdstukken 8 en 9 (de Nederlandse vertaling van hoofdstuk 8) bevatten een samenvatting van de bevindingen en hieruit af te leiden implicaties voor de onderwijspraktijk. De eerste - voor de hand liggende - suggestie betreft het advies om activiteiten te kiezen die de tweede-taalverwerving het meest stimuleren. Uit de studie blijkt dat het werken in een groepje zeer geschikt voor zowel participatie als controle is en dat daarnaast de NT2-lessen van groot belang zijn voor met name de participatie. Damhuis pleit dan ook voor meer speciale NT2-lessen. Het vrije kringgesprek blijkt niet erg stimulerend voor de verwerving van produktieve taalvaardigheid, het is de activiteit waarin de leerlingen de minste initiatieven nemen. De tweede suggestie betreft het aanpassen van gangbare interacties, zodat ze gunstiger worden voor tweede-taalverwerving. Leerkrachten zouden bekend gemaakt moeten worden met de theoretische uitgangspunten van participatie en controle en ze zouden manieren moeten krijgen om deze in handelen te vertalen. Voorbeelden van dergelijke handelingen zijn bijvoorbeeld het geven van meer produktiegelegenheden aan anderstalige kinderen door ze meer tijd te geven om van de aangeboden gelegenheid tot spreken gebruik te maken, het verminderen van het aantal vragen en het vermeerderen van het aantal luisterresponsen om zo het beurtwisselingsinitiatief bij het anderstalige kind te laten, of proberen om de keuze van het onderwerp aan het anderstalige

kind over te laten. Al de geschetste handelingen zijn gericht op het verhogen van participatie en controle van het anderstalige kind. Dit streven is gerechtvaardigd: uit de resultaten van het onderzoek blijkt duidelijk dat de controlebalans vaak te ver is doorgeslagen in de richting van de leerkracht; zelfs bij het werken in een groepje, een activiteit waar de leerkracht in principe slechts zijdelings bij is betrokken, heeft deze vaak nog (gedeeltelijk) controle. Helaas wordt er op deze plaats geen nuancering aangebracht met betrekking tot de mate van controle door de leerlingen: een te veel aan controle van hun kant kan tot vermijdingsstrategieën leiden, hetgeen hun tweede-taalverwerving juist minder gunstig zou beïnvloeden.

Interaction and Second Language Acquisition is een helder geschreven, goed gestructureerd en prettig leesbaar boek, alhoewel de mate van herhaling soms wat irritant werkt. Zo wordt bijvoorbeeld de in het eerste en het afsluitende hoofdstuk terecht opgenomen opmerking dat het onderzoek zich beperkt tot produktie, maar dat het belang van receptie en de rol van input en feedback bij tweede-taalverwerving worden erkend, bij elke dimensie opnieuw herhaald (p. 73, 144 en 226), zodat hij in totaal vijf keer is opgenomen.

Het onderwerp van *Interaction and Second Language Acquisition* is voor zowel onderzoekers als leerkrachten van belang. Toch zou ik het boek met name aan de eerste groep willen aanraden. Het onderzoek is gedegen opgezet, er is veel aandacht besteed aan de kwaliteit van de data en de validiteit en de betrouwbaarheid van de operationalisaties van participatie en controle. De gedetailleerde beschrijving zal zeker van nut zijn bij eventueel vervolgonderzoek. Hierbij moet wel aangetekend worden dat de taxonomieën mijns inziens niet eenvoudig te hanteren zijn. Ook Damhuis merkt op dat de gebruikte maten en taxonomieën pas na een "heel intensieve training en oefening" (p. 277) met een voldoende inter-codeurs betrouwbaarheid konden worden gebruikt. Bij nader onderzoek zouden twee aspecten zeker in aanmerking moeten worden genomen: enerzijds de taalvaardigheid van de anderstalige kinderen (zoals Damhuis zelf ook aangeeft) en anderzijds zou een beeld geschetst moeten worden van de interactiemogelijkheden en dus de leermogelijkheden van de

kinderen buiten de klas, aspecten die Damhuis gezien het tijdsbestek van haar studie helaas niet in het onderzoek heeft kunnen betrekken. Voor mensen uit de onderwijspraktijk vind ik het boek wat minder boeiend. Naast enkele implicaties die worden afgeleid uit de kwalitatieve beschrijvingen van beurtwisselings- en onderwerpsparticipatie en -controle aan de hand van voorbeelden in hoofdstuk 4 en 6, komt de derde onderzoeksvraag eigenlijk alleen in de laatste paragraaf expliciet aan de orde, maar hier worden vrij weinig concrete suggesties gedaan. Damhuis signaleert dit zelf ook, maar voegt daar aan toe: "Maar de operationalisaties en het taalmateriaal uit het huidige onderzoek, maken het mogelijk deze suggesties veel specifiekier uit te werken. Op die manier zou een beknopte cursus in conversatietechnieken voor Participatie en Controle ontwikkeld kunnen worden voor gebruik in de scholing van leerkrachten" (p. 284). Mocht deze cursus op de markt komen, dan is dit wellicht wel een aanrader voor docenten.

M. M. Jagtman

Literatuur

- Lier, L. van (1988). *The classroom and the language learner. Ethnography and second-language classroom research*. London: Longman.
- Keenen, E. O., & Schieffelin, B. B. (1976). Topic as a discourse notion. A study of topic in the conversations of children and adults. In C. N. Li (Ed.), *Subject and topic* (pp. 335-384). New York: Academic Press.

S. Brand-Gruwel

Onderwijs in tekstbegrip. Een onderzoek naar het effect van strategisch lees- en luisteronderwijs bij zwakke lezers

Uitgeverij Tandem Felix, Ubbergen 1995, X + 135 pagina's, f 48,-, ISBN 90 73345 54 5

De laatste jaren wordt veel belang gehecht aan strategisch leesonderwijs, d.w.z. aan leesonderwijs waarin (ook) systematisch aandacht besteed wordt aan het ontwikkelen van strategieën die bij begrijpend en studerend lezen een belangrijke rol spelen. Nu is algemeen bekend dat het onderwijzen van dergelijke leesstrategieën aan leerlingen die veel moeite hebben met technisch lezen, vaak erg moeilijk is, omwille van het decodeerprobleem waarmee zij tijdens het lezen voortdu-

rend worden geconfronteerd. De centrale vraagstelling van dit proefschrift is dan ook of het zin heeft om onderwijs in begripvend leesstrategieën te verstrekken aan zwakke tot zeer zwakke technische lezers, en zo ja hoe dit onderwijs het best georganiseerd wordt. Een oplossing die in dit verband vaak gesuggereerd wordt, is om de beoogde (meta)cognitieve strategieën niet (enkel) via het lezen, maar (ook) via het luisteren te ontwikkelen. Op die manier kan het decodeerprobleem wellicht geheel of gedeeltelijk worden omzeild, zo luidt de redenering.

In Hoofdstuk 1 – het literatuurhoofdstuk – worden de begrippen ‘lezen’ en ‘luisteren’ in een ruimer theoretisch kader geplaatst. Daarna beschrijft de auteur de processen die tijdens het lezen plaatsvinden, evenals de wijze waarop lezers dit proces met behulp van strategieën trachten te reguleren en te evalueren. Ook de relatie tussen begripvend lezen en luisteren komt uitvoerig aan bod. Op grond van de beschikbare empirische gegevens komt de auteur tot de conclusie dat er een heel sterke relatie bestaat tussen luisteren en lezen, en dat er in beide situaties grotendeels beroep wordt gedaan op dezelfde verwerkingsprocessen. Vervolgens besteedt de auteur uitvoerig aandacht aan de problemen die kinderen ondervinden met begripvend lezen en beschrijft zij wat daarvan zoal de oorzaken kunnen zijn. Daarna wordt ingegaan op onderwijs in begripvend lezen en meer bepaald op de vraag hoe tekstbegrip van leerlingen met begripvend leesproblemen via instructie kan worden verbeterd. Daarbij heeft zij vooral oog voor dit soort van onderwijsleeromgevingen waarin de klemtoon sterk komt te liggen op het plannen, reguleren en evalueren van het leesproces. Wat de vormgeving van de onderwijsleeromgeving betreft, worden zowel het ‘directe instructiemodel’ als het ‘reciprocal teaching model’ besproken. Bij de bespreking van dit laatste model komt het werk van Palincsar en Brown zeer uitgebreid aan bod. Daarbij bespreekt de auteur niet enkel de heel bekende studie van 1984, waarin de ‘reciprocal teaching’-methode met succes is aangewend bij het onderwijs aan brugklasleerlingen in vier strategieën die ingezet kunnen worden om het tekstbegrip te verbeteren (‘comprehension forstering’) en te evalueren (‘comprehension monitoring’), nl. (1) ophelderende van onduidelijkheden, (2) vragen stellen, (3) samenvatten en (4) voorspellen van het verdere verloop van de tekst; daarnaast be-

sprekt zij ook nog een minder bekende studie van Palincsar en Brown van 1989, waarin aangetoond is dat het mogelijk is deze vier strategieën door middel van luisteronderwijs ook bij (nog) zwakkere leerlingen te ontwikkelen. Een probleem met deze laatste studie is evenwel dat de resultaten ervan – omwille van een aantal tekorten in de onderzoeksofzet – moeilijk te interpreteren zijn. Dit leidt tot de volgende twee onderzoeksvragen, die door de auteur beantwoord worden via twee gelijkaardige studies.

- (1) Wat is het effect van een lees- en luisterprogramma ‘Elkaar onderwijzen in begripvend lezen en luisteren’ genaamd, waarin kinderen van groep 6 van het basisonderwijs en kinderen van 9 tot 11 jaar van het lom-onderwijs die zwak zijn in decoderen en begripvend lezen én zwak of goed in luisteren zijn, worden getraind in de vier tekstverwerkingsstrategieën.
- (2) Wat is het effect van een luisterprogramma ‘Luisteren met begrip’ genaamd, waarin kinderen van 9, 10 en 11 jaar van het lom-onderwijs die zeer zwak zijn in decoderen, zwak in begripvend lezen én zwak of goed in luisteren zijn, worden getraind in de vier tekstverwerkingsstrategieën.

In hoofdstuk 2 wordt een beschrijving gegeven van alle materialen en meetinstrumenten die ten behoeve van het onderzoek ontwikkeld zijn. Eerst worden van de beide programma’s het doel, de opbouw en de opzet van de lessen met de didactische hulpmiddelen besproken. Uit deze beschrijving blijkt dat de auteur zich heel sterk heeft laten inspireren door het eerder genoemd werk van Palincsar en Brown, maar dit toch heeft aangevuld met enkele elementen uit het directe instructiemodel (nl. voor wat betreft de globale opbouw van het programma en de ordening van de leerstof). Daarna worden de instrumenten beschreven. Dit zijn zowel instrumenten die speciaal voor het onderzoek zijn ontwikkeld als gestandaardiseerde tests. De ontwikkelde meetinstrumenten hadden tot doel na te gaan in hoeverre leerlingen de vier strategieën lezend en luisterend beheersten, en te bepalen hoeveel metacognitieve kennis leerlingen over lezen bezitten. De gestandaardiseerde tests waren eerder bedoeld om het begripvend lezen en luisteren in het algemeen te meten.

In de twee daaropvolgende hoofdstukken worden achtereenvolgens de opzet, de methode

van onderzoek en de resultaten van de twee studies besproken. Uit de eerste studie bleek dat het doel van het programma – de ontwikkeling van de vier strategieën – werd bereikt; ook presteerden de experimentele leerlingen direct na de interventie beter op de gestandaardiseerde begrip- pend leestest en deze resultaten bleken nog te bestaan tijdens de retentiemeting; tenslotte bleken de leerlingen die het programma volgden ook over meer metacognitieve kennis te beschikken dan de kinderen uit de controlegroep. Interessant is tenslotte dat er geen opvallende effectverschillen gevonden werden tussen de kinderen uit het reguliere onderwijs en de lom-kinderen, en ook niet tussen kinderen die goed en zwak waren in begrijpend luisteren. In de tweede studie werd vastgesteld dat de kinderen aanzienlijk beter geworden waren in het luisterend uitvoeren van de vier strategieën, maar voor de rest waren de resultaten nogal teleurstellend. Ook de verwachte effect-verschillen tussen onderscheiden groepen leerlingen bleven uit.

Zoals gewoonlijk in dit type van onderzoek, stelt zich bij de interpretatie het lastige probleem welke aspecten van de programma's verantwoordelijk zijn voor het (succesvolle) effect ervan. In het slothoofdstuk zet de auteur de verschillende factoren nogmaals netjes op een rijtje, maar verder dan de conclusie dat ze waarschijnlijk allemaal (gedeeltelijk) bijgedragen hebben tot de gevonden effecten en dat derhalve verder onderzoek nodig is komt zij niet. Na een aantal beperkingen en knelpunten te hebben opgesomd, geeft de auteur tenslotte een reeks waardevolle suggesties voor dergelijk vervolgonderzoek.

Het proefschrift van Brand-Gruwel is een belangrijk boek voor al wie vanuit wetenschappe-

lijk en/of praktisch oogpunt bezig is met de analyse en bevordering van tekstbegrip. Het literatuurhoofdstuk is van hoge kwaliteit. De relevante literatuur wordt er op overzichtelijke wijze samengebracht, beknopt én kritisch besproken, en uitgewerkt tot een heldere en scherpe vraagstelling. Alleen het onderscheid tussen cognitieve en metacognitieve activiteiten en strategieën – één van de kernproblemen in dit onderzoeksgebied – wordt m.i. door de auteur niet altijd even scherp gemaakt en aangehouden. Ook de empirische hoofdstukken lezen aangenaam en vlot, niet enkel omwille van de vlotte en heldere schrijfstijl maar ook omwille van de optimale afstemming tussen de presentatie van onderzoeksvragen en -hypothesen, onderzoeksopzet, resultaten en conclusies. Natuurlijk vormt het ontbreken van een afdoende verklaring van het (positieve) effect van de beide programma's – zoals gezegd – een belangrijk tekort, maar zolang er nog grote onduidelijkheid bestaat omtrent de haalbaarheid en effectiviteit van strategisch gerichte instructie, blijft de behoefte bestaan aan goed opgezette trainingsstudies zoals die van Brand-Gruwel waarin de aandacht primair uitgaat naar het behalen van een zo groot mogelijk leer- en transfereffect. Om al deze redenen verdient het boek aanbeveling bij al wie rechtstreeks betrokken is bij leesonderwijs aan zwakke lezers. Maar ook onderzoekers en practici die in andere vakgebieden (bijv. wiskunde, zaakvakken...) en op andere leeftijdsniveaus (bijv. secundair onderwijs) werken aan de analyse en ontwikkeling van (meta)cognitieve strategieën, zullen in dit boek heel wat waardevolle en inspirerende inzichten en ideeën vinden.

L. Verschaffel

Inhoudsopgave

Pedagogische Studiën Jaargang 73, nummer 6/1996

R.H. Hofman en C.J.W. Meijer

Weer Samen Naar School: uitkomsten van onderzoek en toekomstperspectief. Introductie op het thema 402

C.J.W. Meijer

Weer Samen Naar School in cijfers 404

R.H. Hofman en E.B. Vonkeman

Onderwijsmethoden en adaptief onderwijs 411

A.A.M. Houtveen, N. Booij, R. de Jong en W.J.C.M. van de Grift

Adaptief onderwijs en leerlingresultaten 422

B. Petersen en D. Oudenhoven

Waar komen de regionale verschillen in verwijzing vandaan? 434

C.J.W. Meijer

Weer Samen Naar School: terugblik en vooruitzicht 441

M. Verkuyten, B. Kinket en Ch. van der Wielen

Oudere kinderen en hun begrip van discriminatie 447

J.H. Walma van der Molen en T.H.A. van der Voort

Leereffecten van Jeugdjournaal en gedrukt nieuws bij kinderen: Een mediumvergelijkend experiment 459

Boekbesprekingen 474

Mededelingen 478

BIBLIOTHEEK DER
RIJKSUNIVERSITEIT
UTRECHT

Weer Samen Naar School:

uitkomsten van onderzoek en toekomstperspectief

Introductie op het thema

R. H. Hofman en C. J. W. Meijer

Weer Samen Naar School

Al langere tijd is in Nederland de groei van het speciaal onderwijs een bron van aanhoudende zorg. De overheid heeft door de jaren heen dan ook verschillende maatregelen genomen en experimenten gestimuleerd om de uitstroom uit het regulier onderwijs te verminderen. Sinds het schooljaar 1992/1993 zijn scholen voor regulier en speciaal basisonderwijs opgegaan in samenwerkingsverbanden in een poging de groei van het speciaal onderwijs een halt toe te roepen. Dit beleid staat bekend onder de naam 'Weer Samen Naar School' (WSNS). In 1995 zijn er zo'n kleine 300 samenwerkingsverbanden werkzaam van gemiddeld 30 scholen voor regulier onderwijs en twee scholen voor speciaal onderwijs. Die samenwerkingsverbanden herbergen in totaal bijna 1,5 miljoen leerlingen. Daarin wordt onder andere getracht de samenwerking tussen speciaal en regulier onderwijs inhoudelijk vorm te geven, een systeem van leerlingzorg op schoolniveau te creëren en adaptief onderwijs in de klas tot stand te brengen.

In hun analytische beleidsevaluatie hebben Meijer, Meijnen en Scheerens (1993) de kans op succes van het WSNS-beleid direct verbonden aan de mate waarin het onderwijs er in slaagt om te gaan met verschillen tussen leerlingen. Adaptief onderwijs wordt daarbij gezien als de belangrijkste inhoudelijke voorwaarde voor het adequaat opvangen van leerlingen met specifieke pedagogisch-didactische behoeften. In het WSNS-beleid gaat het dus om een betere toerusting van het basisonderwijs door middel van vergroting van deskundigheid en het bieden van faciliteiten om conform de doelstellingen van de Wet op het Basisonderwijs (zorgverbreding, ononderbroken ontwikkeling) eerder zorg te bieden die meer flexibel is georganiseerd, hetgeen moet

leiden tot een vermindering van de schooluitval (Meijer, Meijnen & Scheerens, 1993).

Adaptief onderwijs houdt rekening met verschillen tussen leerlingen en stemt het onderwijsaanbod daarop af. Het vertrekpunt van adaptief onderwijs is het metterdaad vaststellen van de verschillen tussen leerlingen door de leerkracht. Het leerkrachtenteam dient geen negatieve consequenties te verbinden aan de geconstateerde verschillen en daarenboven te streven naar het behalen van een zo hoog mogelijk prestatieniveau voor *alle* leerlingen, risicoleerlingen, modale en ook (hoog-) begaafde leerlingen (Meijer, Meijnen & Scheerens, 1993). In het perspectief van Weer Samen Naar School is de leerkracht, bij het afstemmen van het onderwijs op verschillen tussen leerlingen, niet de alleen-verantwoordelijke, maar er kan ook hulp van deskundigen of andere leerkrachten worden ingeroepen.

Evaluatieprogramma WSNS

Toen in 1992 het Weer Samen Naar School-beleid (WSNS) werd ingezet, werd tegelijkertijd een wetenschappelijk gefundeerd evaluatieprogramma ontworpen. De Stuurgroep Evaluatie WSNS heeft in het Uitvoeringsplan (Bosker & Meijer, 1993) een aantal deelonderzoeken uitgezet en er worden op dit moment verschillende studies verricht die vanuit een bepaalde invalshoek kennis opleveren over het functioneren van de samenwerkingsverbanden WSNS. Vier vragen dienen door de Stuurgroep te worden beantwoord, namelijk:

- 1 of risicoleerlingen in voldoende mate kunnen worden opgevangen met de voorzieningen die de samenwerkingsverbanden bieden,
- 2 of en hoe het streefbeeld in de praktijk realiseerbaar is,

- 3 de mogelijkheden van verevening te onderzoeken en
- 4 jaarlijks nagaan of stabilisatie van leerling-aantallen plaatsvindt.

Om vast te kunnen stellen in hoeverre scholen in staat zijn leerlingen met speciale zorg op te kunnen vangen, moet men zowel naar het samenwerkingsverband, de school, de leerkrachten als de leermiddelen kijken. Inmiddels is een instrument ontwikkeld om vast te stellen in hoeverre scholen adaptief onderwijs kunnen aanbieden (Houtveen & Booij, 1994). Momenteel wordt in grootschalig onderzoek vastgesteld in welke mate die situatie in het Nederlandse basisonderwijs van de grond komt. Verder worden risicoleerlingen gevolgd en de kwaliteit van de zorg die zij ontvangen nagegaan. Ook op het niveau van de samenwerkingsverbanden wordt onderzoek verricht. Naast een jaarlijkse registratie en beschrijving van verwijzingscijfers en aanverwante gegevens van de samenwerkingsverbanden door het IVA (de zgn. IVA-monitor), wordt onderzoek verricht naar de organisatie en effecten van de samenwerkingsverbanden. Ook worden enkele voorlopende scholen uitvoerig beschreven. Tenslotte zijn veel onderzoeksinspanningen gericht op het verkrijgen van informatie over de rol van ouders, over de stand van zaken van ambulante begeleiding, etc.

Enkele van de genoemde onderzoeken, geëntameerd door de Stuurgroep Evaluatie WSNS, hebben reeds resultaten opgeleverd. Daarom lijkt het een goede zaak om de bevindingen van die studies, maar ook daaraan verwant onderzoek op samenhangende wijze te presenteren aan belangstellenden die vanuit een meer wetenschappelijke achtergrond geïnteresseerd zijn in het functioneren van grootschalige experimenten zoals WSNS.

Dit themanummer start met een bijdrage van Meijer, coördinator van de Stuurgroep Evaluatie Weer Samen Naar School, waarin hij de actuele stand van zaken van het beleidsproces met cijfers onderbouwt. Daarna volgt de bijdrage van Hofman en Vonkeman waarin zij beschrijven in hoeverre gangbare en nieuwe onderwijsmethoden voor rekenen, spelling en lezen een geschikte conditie vormen voor het realiseren van adaptief onderwijs in de klas. De bijdrage van Houtveen, Booij, De Jong en Van

de Grift gaat in op de relatie tussen adaptief onderwijs en leerlingresultaten. In de bijdrage van Petersen en Oudenhoven wordt nagegaan welke regionale factoren invloed uitoefenen op (verschillen in) verwijzing van leerlingen naar het speciaal onderwijs. Het themanummer wordt afgesloten met een samenvatting en discussie door de coördinator van de Stuurgroep Evaluatie Weer Samen Naar School. Meijer geeft een overzicht van de stand van zaken, vat de belangrijkste resultaten nog eens samen en belicht deze in het perspectief van het toekomstig WSNS-beleid (vgl. ook Meijer, 1995). Met name de voorwaarden voor succes enerzijds en de implicaties voor het vigerende WSNS-beleid anderzijds, komen daarbij voor het voetlicht. Bovendien schenkt hij aandacht aan relevante buitenlandse bevindingen voor de integratie van regulier en speciaal onderwijs die toepasbaar zijn in de Nederlandse situatie.

Literatuur

- Bosker, R., & Meijer, C. (1993). *Uitvoeringsplan voor de Evaluatie van het Weer Samen Naar School-beleid*. De Lier: Academisch Boeken Centrum.
- Houtveen, A. A. M., & Booij, N. (1994). *Het meten van integrale leerlingzorg. Adaptief onderwijs en schoolontwikkeling*. Utrecht: ISOR, Universiteit Utrecht.
- Meijer, C. J. W. (1995). *Halverwege: van startwet naar streefbeeld*. Onderzoek in opdracht van de Stuurgroep Evaluatie Weer Samen Naar School. De Lier: Academisch Boeken Centrum.
- Meijer, C., Meijnen, W., & Scheerens, J. (1993). *Over wegen en schatten en sturen*. Analytische beleidsevaluatie Weer Samen naar School. De Lier: Academisch Boeken Centrum.

C. J. W. Meijer

Samenvatting

Het Weer Samen Naar School-beleid lijkt nog niet veel vruchten te hebben afgeworpen, tenminste als we deze afmeten aan de absolute groei van de deelname aan de groep-1 scholen (LOM, MLK en IOBK). Voorzichtigheid is echter geboden. Nader onderzoek naar de stroomgegevens en naar de bijdrage van effecten zoals ontwikkelingen in de verblijfsduur, zal meer licht op de kwantitatieve ontwikkelingen kunnen bieden. De samenwerkingsverbanden laten grote verschillen zien: er zijn enkele zeer grote, maar ook enkele zeer kleine verbanden. Ook de regionale verschillen in deelnamecijfers zijn aanzienlijk. Urbanisatiegraad en het aantal migrantenkinderen zijn hierbij als (deels overlappende) factoren voor verantwoordelijk. Met name het voorzieningenniveau wordt aangedragen als belangrijke determinant van verwijzing naar het speciaal onderwijs.

1 Inleiding

Een overzicht van de recente stand van zaken van het Weer Samen Naar School-proces kan niet zonder een beschouwing over de kwantitatieve ontwikkelingen. Niet in de laatste plaats omdat het stabilisatievraagstuk zo'n centrale rol inneemt. Bij het bereiken van het zogenaamde *3 x Akkoord* (Ministerie van O&W., 1991) hebben de besturen-, vak- en ouderorganisaties stabilisatie immers als hun 'inspanningsverplichting' gedefinieerd. In de volgende paragrafen wordt daarvan de balans opgemaakt. Hierbij wordt niet alleen ingegaan op de ontwikkelingen in de zogenaamde groep-1 scholen (leer- en opvoedingsmoeilijkheden - LOM -, moeilijk lerende kinderen - MLK - en de aan deze typen verbonden afdelingen voor in hun ontwikkeling bedreigde kleuters - IOBK-), maar komen ook ontwikkelingen in het overige speciaal onderwijs (SO) en voortgezet speciaal onderwijs (VSO) aan de

orde. Eerst wordt kort de stand van zaken in de samenwerkingsverbanden besproken (2). Daarna wordt ingegaan op het stabilisatievraagstuk (3). Vervolgens worden de beschikbare stroomgegevens geanalyseerd (4). De bijdrage wordt afgesloten met een korte beschouwing over een aantal andere aspecten die van belang zijn voor de WSNS-discussie, zoals de regionale verschillen in de deelname aan het speciaal onderwijs en de deelname van allochtonen (5). Het hier gepresenteerde cijfermateriaal is grotendeels gebaseerd op de tussenrapportage van de Stuurgroep Evaluatie WSNS *Halverwege* (Meijer, 1995).

2 De samenwerkingsverbanden in cijfers

Er is in korte tijd een landelijk dekkend netwerk van bijna 300 samenwerkingsverbanden (SWV's) tot stand gebracht. Vrijwel alle basisscholen nemen nu deel aan een SWV. Gemiddeld bestaat een SWV uit circa 30 scholen, waarvan er twee behoren tot het speciaal onderwijs (groep 1). Een SWV heeft gemiddeld 5000 leerlingen, waarvan bijna 200 leerlingen in het speciaal onderwijs (IVA, 1994a). Er zijn grote verschillen in omvang (zie Boef-van der Meulen, Schüssler & Karsten, 1993). Er zijn verbanden met minder dan 10 scholen (circa 20 SWV's) en er zijn verbanden met meer dan 50 scholen (circa 50 SWV's). Het kleinste verband telt één basisschool en één school voor speciaal onderwijs, de grootste meer dan 140 basisscholen en circa 10 scholen voor speciaal onderwijs. De range in aantallen leerlingen loopt van 200 tot ver over de 20.000! In de praktijk zijn de grotere verbanden veelal weer onderverdeeld in kleinere subverbanden.

Ruim tweederde van de SWV's is interzuilair en de rest is zuilair georganiseerd. In zo'n 80 verbanden participeert één school voor speciaal onderwijs, in de overige SWV's zijn meer scholen voor speciaal onderwijs opgenomen.

In bijna de helft van de 290 SWV's participeert een LOM- of MLK-school met een IOBK-afdeling. De andere SWV's hebben dus niet een dergelijke voorziening. Er zijn ook enkele andere typen speciaal onderwijs betrokken bij de SWV's: vijf scholen voor zeer moeilijk opvoedbare kinderen en twee scholen voor langdurig zieke kinderen nemen aan de SWV's deel.

3 Stabilisatie

De inspanningsverplichting om de deelnamecijfers van groep-1 scholen te stabiliseren is een in het oog lopende afspraak uit *3 x Akkoord*. De Minister verstaat onder stabilisatie (Memorie van Antwoord, 23486, nr. 5, 14-2-1994): "Van stabilisatie is sprake als voldaan is aan de volgende voorwaarde: Het aantal leerlingen in het LOM, MLK en IOBK onderwijs overschrijdt op de peildata (jaarlijks 1 oktober) de stabilisatienorm (aantal leerlingen per 16-1-92) niet. (...) Hoewel het WSNS-beleid zich beperkt tot de scholen van groep 1, houdt het akkoord expliciet rekening met de ontwikkelingen in groep 2/3 en strekt de inspanningsverplichting van de organisaties zich ook uit tot deze sector voor zover het om weglek-effecten gaat. Het VSO valt buiten het bereik van deze afspraken" (Tweede Kamer der Staten Generaal, 1994a). Het akkoord heeft dus een breder perspectief dan de groep-1 scholen en als zodanig zullen we ook de ontwikkelingen in de andere vormen van speciaal onderwijs aan de orde stellen. Overigens zullen we ook kort ingaan op het voortgezet speciaal onderwijs, hoewel deze sector buiten het bereik van het akkoord ligt.

Bij stabilisatie gaat het formeel om de vergelijking van de deelnamecijfers in de groep-1 scholen tussen het referentie-metmoment van 16-1-1992 en de jaarlijkse oktoberpeiling. Op zich is dit een merkwaardige afspraak. Een peiling in januari laat zich immers moeilijk vergelijken met een meetmoment in oktober. Er kan namelijk sprake zijn van seizoensinvloeden. In het volgende zullen wij ons daarom beperken tot vergelijkingen over dezelfde meetmomenten.

Uit de peilingen in oktober blijkt een lande-

lijke stijging. Tussen oktober 1991 en oktober 1993 steeg het aantal leerlingen in het onderwijs op groep-1 scholen met bijna 1000 leerlingen. Uit die cijfers blijkt dat met name het IOBK en het MLK zijn gestegen. Het LOM groeide licht (zie Tweede Kamer der Staten Generaal, 1994b). Strikt genomen is dus nog niet aan de stabilisatievoorwaarde voldaan: het aantal leerlingen in de groep-1 scholen is immers gestegen. Indien we echter rekening houden met de toename van de basisgeneratie (alle 4-12 jarige kinderen), dan kunnen we vaststellen dat er geen sprake is van groei maar van stabilisatie: de bevolkingsstatistieken van januari 1994 laten een toename zien van het aantal 4-12 jarigen van ruim 25.000 ten opzichte van 1992! Daarmee rekening houdend blijft de procentuele deelname aan het speciaal onderwijs (groep 1) op de aangegeven meetmomenten tamelijk stabiel.

De groep-1 scholen laten in absolute aantallen dus een lichte groei zien. Ook het totaal aantal leerlingen in het totale (voortgezet) speciaal onderwijs is toegenomen: tussen oktober 1992 en oktober 1993 nam het (V)SO toe van ruim 111.000 leerlingen tot ruim 114.000 leerlingen; een toename van ruim 3000 leerlingen, ofwel een stijging van 2.8% (zie Bijlage 1 bij de brief van de Minister aan de Tweede Kamer, 22 maart 1994). Het speciaal onderwijs steeg met bijna 1500 leerlingen en het voortgezet speciaal onderwijs met ruim 1600 leerlingen.

Voor wat betreft het speciaal onderwijs zijn de grootste stijgers: scholen voor kinderen met ernstige spraakmoeilijkheden (ESM), in hun ontwikkeling bedreigde kleuters (IOBK), zeer moeilijk lerende kinderen (ZMLK), meervoudig gehandicapten (MG) en moeilijk lerende kinderen (MLK):

ESM:	+ 144 leerlingen	=	6.26%
IOBK:	+ 175 leerlingen	=	5.88%
ZMLK:	+ 247 leerlingen	=	4.31%
MG:	+ 124 leerlingen	=	4.12%
MLK:	+ 511 leerlingen	=	2.21%

(N.B. alleen typen met een groei boven de 100 leerlingen zijn hier gepresenteerd)

In het voortgezet speciaal onderwijs is vooral sprake van een stijging bij moeilijk lerende kinderen (MLK), zeer moeilijk lerende kinderen (ZMLK), zeer moeilijk opvoedbare kinderen (ZMOK) en kinderen met leer- en opvoedings-

moelijkheden (LOM):

MLK:	+ 973 leerlingen	=	7.66%
ZMLK:	+ 151 leerlingen	=	5.39%
ZMOK:	+ 181 leerlingen	=	4.82%
LOM:	+ 223 leerlingen	=	2.04%

Deze gegevens maken duidelijk dat de ontwikkeling in de groep-1 scholen in het SO maar een deel van het (groei-)probleem vormt. Met name in het voortgezet speciaal onderwijs is sprake van een sterke stijging van het aantal leerlingen in de hier genoemde schoolsoorten. Het VSO-LOM en het VSO-MLK nemen samen 1200 leerlingen voor hun rekening van de totale groei van het voortgezet speciaal onderwijs van ruim 1600 leerlingen.

De vraag is of hierbij sprake is van autonome ontwikkelingen of dat er een samenhang is met WSNS-beleid. Zo zou de groei in het voortgezet speciaal onderwijs mede verklaard kunnen worden door een eerdere instroom vanuit de bovenbouw van het speciaal onderwijs, leidend tot een kortere verblijfsduur in het speciaal onderwijs. Maar het kan ook verklaard worden uit een algemene verlenging van de verblijfsduur in het voortgezet speciaal onderwijs en/of uit een toegenomen uitstroom uit het reguliere voortgezet onderwijs.

Ook op het niveau van de overige schoolsoorten voor speciaal onderwijs kan sprake zijn van met WSNS samenhangende effecten. De groei van het ESM en het ZMLK zou verklaard kunnen worden doordat een deel van de LOM/MLK/IOBK-populaties nu bij het ESM en ZMLK wordt aangemeld. Voor het ZMLK zijn daar inderdaad indicaties voor: de ontwikkeling in dit schooltype werd de laatste tien jaar gekenmerkt door een opvallende stabiliteit. Tussen 1983 en 1990 schommelde de omvang zo rond de 5200 à 5300 leerlingen. De bijna 6000 leerlingen die nu naar het ZMLK gaan, vormen op die trend een afwijking. De leerlingen aan de 'onderkant' van het MLK zouden in die hypothese dus eerder een overstap maken naar het ZMLK-onderwijs.

Ook wat het ESM betreft zijn al eerder indicaties verkregen die wijzen op een overlap met de LOM/MLK/IOBK-problematiek. De ARBO concludeerde in haar advies 'Omgaan met verschillen' dat het onderscheid tussen LOM/MLK en ESM niet van essentiële aard is (ARBO, 1991). De verschillen zijn slechts gra-

deel. De Raad ging na een uitgebreide analyse op bronnen en vele consultaties zelfs zover te adviseren niet langer een speciale school voor deze leerlingen in stand te houden, maar ze te integreren in de door de Raad voorgestelde nieuwe speciale school of in de verbrede basis-school.

4 Een analyse van de stromen tussen regulier en speciaal onderwijs

Over de totale groep-1 scholen heen blijkt dat van een WSNS-effect in termen van een stabilisatie van de absolute aantallen nog geen sprake is. Hiervoor werd reeds gememoreerd dat de oktobervergelijkingen een groei lieten zien. Deze groei is in principe de resultante van een aantal factoren: instroom uit het basisonderwijs, terugstroom naar het basisonderwijs, uitstroom naar het voortgezet speciaal onderwijs, naar het reguliere voortgezet onderwijs, naar andere typen speciaal onderwijs, maar ook van verblijfsduurveranderingen. Een analyse van het IVA (van de cijfers van oktober 1992 en 1993) laat zien dat de totale instroom uit het basisonderwijs is gestegen met bijna 700 leerlingen en dat de terugstroom is gedaald met ruim 100 leerlingen. Verder stijgt de uitstroom uit de groep-1 scholen naar het voortgezet speciaal onderwijs met ruim 200 leerlingen, daalt de uitstroom naar het reguliere voortgezet onderwijs met eveneens ruim 100 leerlingen en stijgt de uitstroom naar het overige speciaal onderwijs met dezelfde omvang. Samenvattend: geen van deze ontwikkelingen is conform de doelstellingen van WSNS. Indien we corrigeren voor de groei van de 4-12 jarigen in de basisgeneratie blijkt echter sprake te zijn van een stabilisatie.

Recent heeft het IVA (1994b) een onderzoek verricht naar de ontwikkelingen zoals die zich voordoen met betrekking tot het onderwijs aan in hun ontwikkeling bedreigde kleuters (IOBK). Op grond van de gegevens zoals die door CBS/Cfi zijn verzameld over de periode 1991-1993 blijkt het volgende. Het IOBK groeide in die periode met zo'n 250 leerlingen waarvoor vooral het IOBK-LOM verantwoordelijk is. De instroom uit het basisonderwijs is gegroeid met 170 leerlingen. Vooral uit groep 1

van het basisonderwijs zijn meer leerlingen verwezen naar het IOBK. Ook de 'onderinstroom' in het IOBK groeit. In totaal is over genoemde periode een toename van instroom uit de 'niet-onderwijssector' waar te nemen met bijna 200 leerlingen. Vooral uit medische kinderdagverblijven en uit peuterspeelzalen worden steeds meer kinderen verwezen (met name naar LOM-IOBK).

Verder constateert het IVA een duidelijk 'voorsorteer-effect'. Er wordt vooral uit het IOBK uitgestroomd naar de 'eigen' school voor speciaal onderwijs (waaraan het IOBK als afdeling is verbonden). Ongeveer 50% van de uitstroom uit IOBK-LOM gaat naar het LOM, 60% van IOBK-MLK naar het eigen MLK en circa 40% van het IOBK-ZMOK naar het ZMOK.

De terugstroom vanuit het IOBK naar het basisonderwijs is gering en lijkt verder terug te lopen. Met name de terugstroom uit het LOM-IOBK is teruggelopen. In 1993 zijn maar 49 leerlingen teruggeplaatst uit het LOM-IOBK, terwijl in de jaren daarvoor nog bijna het dubbele aantal leerlingen werd teruggeplaatst.

Ook meer in het algemeen is te zien dat van terugstroom niet veel terecht komt. CBS-gegevens wijzen er op dat jaarlijks ongeveer 400 leerlingen uit het LOM- en MLK-onderwijs terugstromen naar het basisonderwijs, een terugstroom die tamelijk stabiel lijkt, hoewel de laatste gegevens op een daling wijzen tussen oktober 1992 en oktober 1993 met ongeveer 30% (IVA, 1994a).

Wat betreft de verwijzingscijfers doen zich ook een aantal interessante ontwikkelingen voor. Uit gegevens van het IVA (1994a), gebaseerd op de terugrapportages vanuit de samenwerkingsverbanden, is een trend van toename in verwijzingen waar te nemen. In 1989 werd nog 0.81% van de basisschoolpopulatie verwezen naar het speciaal onderwijs. In schooljaar 1992/1993 is dit gestegen naar 0.92%. Meijer (1988) rapporteerde over het schooljaar 1984/1985 nog een percentage van 0.74% voor de groepen 3 tot en met 8 van het basisonderwijs. Indien daar het toenmalige kleuteronderwijs bij wordt betrokken, dan ligt dat percentage in de buurt van de 0.80% voor dat schooljaar. Jaarlijks worden dus nog steeds meer leerlingen naar het speciaal onderwijs

verwezen. Het accent ligt hierbij op verwijzingen uit de groepen 2, 3 en 4 van het basisonderwijs. Zie hiervoor Tabel 1.

Tabel 1

Instroom in het SO vanuit het BO; Procentuele verdeling over de verschillende groepen in het basisonderwijs. Bron: CBS, 1991

	MLK	LOM	Totaal SO
groep 1	4%	8%	8%
groep 2	28%	17%	21%
groep 3	27%	19%	21%
groep 4	18%	21%	19%
groep 5	10%	13%	11%
groep 6	4%	5%	5%
groep 7	3%	3%	3%
groep 8	6%	15%	11%

Ruim 60% van de verwijzingen naar het speciaal onderwijs vindt dus plaats in de groepen 2, 3 en 4. Verwijzingen naar het LOM-onderwijs vinden vooral plaats in de groepen 2, 3, 4 en 8. Voor het MLK-onderwijs geldt een concentratie in de groepen 2, 3 en 4.

5 Regionale verschillen

De regionale verschillen in deelnamecijfers staan de laatste tijd sterk in de belangstelling. In de WSNS-discussie speelt het vereveningsvraagstuk immers een essentiële rol. Recente beleidsvoornemens duiden op de toepassing van het vereveningsprincipe in de nabije toekomst. Daarbij zullen alle regio's in principe op eenzelfde hoeveelheid 'zorgmiddelen' kunnen rekenen. De middelen worden dan niet meer gebaseerd op het aantal leerlingen dat deelneemt aan het speciaal onderwijs, maar op het totaal aantal leerlingen in de regio. De vraag is hierbij aan de orde hoe de sterke regionale verschillen in deelname zijn te duiden. Zijn die nu het gevolg van verschillen in leerlingenpopulatie of weerspiegelen de regionale verschillen veeleer - al of niet historisch gegroeide - verschillen met betrekking tot de (opvatting over de) vormgeving van de speciale zorg?

Bij de interpretatie van de regionale verschillen moet allereerst rekening worden gehouden met de invloed die uitgaat van de deelname van allochtone leerlingen. De toename van de deelname van allochtonen aan het speciaal onderwijs is een bekend gegeven (zie Doornbos & Stevens, 1987 en Wijnstra, 1987).

Gegevens uit het Onderwijsverslag over 1992 laten zien dat de proportie allochtone leerlingen in het speciaal onderwijs toe is genomen van 9% in 1986 naar 13% in 1992. In absolute cijfers gaat het om een stijging met bijna 4000 leerlingen in zeven jaar (van ruim 6000 in 1986 naar ruim 10.000 leerlingen in oktober 1992). De toename doet zich vooral gelden in het MLK-onderwijs; daar neemt de proportie toe van 12% naar 18%. De groei van het speciaal onderwijs is daarmee voor een belangrijk deel te verklaren door de toename van allochtone leerlingen. In dezelfde vergelijkingsperiode als die waarin de toename van 4000 allochtonen in het speciaal onderwijs is waargenomen, neemt de omvang van het totale speciaal onderwijs met bijna 7000 leerlingen toe. De groei is dus voor meer dan de helft te verklaren door de toename van het aantal allochtonen!

De Keyzer (1994) heeft de situatie in Utrecht aan een nadere beschouwing onderworpen. Zij laat ook zien dat een groot deel van de instroom in het speciaal onderwijs voor rekening komt van allochtone kinderen. Van de totale instroom in het (V)SO is 35% allochtoon. Dit komt overigens in Utrecht aardig overeen met het percentage allochtonen in het reguliere onderwijs. Marokkaanse kinderen gaan met name naar het MLK en Surinaamse kinderen stromen vooral naar het LOM uit. Turkse kinderen stromen gelijkmatiger naar LOM en MLK uit, maar er gaan ook relatief veel Turkse kinderen naar het onderwijs voor slechtfunctionerende kinderen en voor kinderen met ernstige spraakmoeilijkheden.

Boef-van der Meulen et al. (1993) rapporteren over regionale verschillen op het niveau van de regio's van schoolbegeleidingsdiensten (SBD's). Hun analyse van de situatie in de SBD's laat zien dat de verhouding tussen het aantal leerlingen in het speciaal onderwijs en het aantal leerlingen in het basisonderwijs op SBD-niveau bepaald niet evenwichtig is. De gemiddelde SBD-regio laat een verhouding zien van 1 : 24 tussen het aantal leerlingen in het speciaal onderwijs en het aantal leerlingen in het basisonderwijs. Er zijn echter SBD-regio's waar de verhouding 1 : 10 is en in andere wordt 1 : 112 (!) geconstateerd. Rekening houdende met het feit dat sommige SO-scholen meerdere SBD-gebieden 'bedienen', kan worden geconstateerd dat de deelnamepercentages

variëren tussen ruim 1% en 6% (Meijer, 1995).

Boef-van der Meulen et al. (1993) geven aan dat de verschillen ook op SWV-niveau ver uiteenlopen. Er zijn SWV's waar op elke leerling in het speciaal onderwijs minder dan 10 basisonderwijs-leerlingen worden geteld en andere SWV's waar dat meer dan 40 zijn. Ook in de verwijzingscijfers zijn deze verschillen tussen SWV's aan te treffen. Het IVA laat zien dat er SWV's zijn met een verwijzingspercentage lager dan 0.5%, bij een landelijk gemiddelde van bijna 1%. Aan de andere kant zijn er SWV's met meer dan 2% verwijzingen per jaar.

Een andere studie van het Sociaal en Cultureel Planbureau (Blank, Boef-van der Meulen, Bronneman-Helmers, Herweijer, Kuhry & Schreurs, 1990) wijst ook op de grote regionale verschillen. Uitgaande van een nog hoger aggregatieniveau (dan SWV's of SBD's), namelijk de 25 onderwijsgebieden die het SCP en het CBS onderscheiden, zijn er gebieden met een deelname aan het speciaal onderwijs die twee keer zo hoog is als in andere gebieden! De gebieden met de hoogste deelnamecijfers zijn verspreid: Oost-Groningen, Zuidelijke Achterhoek, Noord-Brabant en Zuid-Limburg. Ook de 'lage' gebieden liggen verspreid over het land: Friesland, Zwolle-Meppel en Dordrecht-Gorinchem. Oost-Groningen heeft 45 LOM- en MLK-leerlingen per 1000 leerlingen en Friesland slechts 21.

Er zijn gebieden waar de verhouding tussen LOM- en MLK-deelname 3 : 1 bedraagt terwijl de landelijke verhouding 1.5 : 1 is. Maar er zijn ook gebieden waar meer MLK-leerlingen aan het onderwijs deelnemen dan LOM-leerlingen (5 : 3). Het SCP verwijst hierbij naar de invloed van het voorzieningenniveau (zie ook de bijdrage van Petersen en Oudenhoven in dit nummer). In gebieden met een laag voorzieningenniveau komen nogal eens leerlingen in het MLK-onderwijs terecht die in gebieden met een hoog voorzieningenniveau in een LOM-school worden geplaatst.

Ook wijzen Blank et al. erop dat in gemeenten met voorzieningen voor speciaal onderwijs de deelname aan het LOM en MLK bijna anderhalf maal hoger is dan in gemeenten zonder voorzieningen. Bovendien zijn de voorzieningen uitwisselbaar. Staat er alleen een LOM-school of alleen een MLK-school, dan is de

deelname navenant: in acht gemeenten met LOM- maar geen MLK-onderwijs is de ratio LOM : MLK 6.3 : 2.5, terwijl in vier gemeenten met wel een MLK-school, maar geen LOM-school, de verhouding LOM : MLK 1.1 : 1.5 is.

Blank et al. besluiten met te stellen dat de gesignaleerde verschillen tussen gemeenten niet terug te voeren zijn op de omvang van de problemen of het soort problemen waarmee leerlingen te kampen hebben. De deelnameverschillen worden door Blank et al. verklaard door de aanwezigheid van voorzieningen. Zij spreken zelfs van een 'sterke beïnvloeding' van de deelnamecijfers door het voorzieningenniveau. Een conclusie die overeenkomt met de in dit nummer gerapporteerde bevindingen van Petersen en Oudenhoven.

6 Tot slot

WSNS lijkt nog niet erg succesvol. Daarop lijken de hier vermelde deelnamecijfers te wijzen. In absolute zin is sprake van een stijging van de deelnamecijfers. De stijging van de basissgeneratie is echter ook aanzienlijk. Een correctie daarvoor resulteert in een stabilisatie van de deelname. Niettemin verwijst de stabilisatie-afpraak slechts naar ontwikkelingen in absolute aantallen. En dan is er inderdaad sprake van groei. Men moet er zich verder van bewust zijn dat allerlei effecten van invloed kunnen zijn op deelnamecijfers. Verblijfsduureffecten, de invloed van stromen binnen en tussen schoolsoorten alsmede demografische ontwikkelingen, zijn enkele belangrijke factoren die in ogenschouw moeten worden genomen. Bovendien is er sprake van factoren die beïnvloed kunnen worden door WSNS en andere groei-determinanten. WSNS is in eerste instantie gericht op een vermindering van de uitstroom uit het basisonderwijs en/of op een verhoging van de terugstroom vanuit groep 1 scholen naar het basisonderwijs. Het is niet gericht op een verandering in verblijfsduur, onderinstroom in het speciaal onderwijs of instroom vanuit andere vormen van speciaal onderwijs in groep 1 scholen. Bij het bepalen van de mate van succes van WSNS is dit een belangrijk gegeven. De Stuurgroep Evaluatie WSNS heeft nader onderzoek uitgezet naar de relatieve invloed van factoren als verblijfsduur op de groei van het speciale

stelsel. Dit onderzoek zal in 1996 worden afgerond.

Literatuur

- ARBO (1991). *Omgaan met verschillen*. Advies over de ontwikkeling van de groep twee-scholen in het speciaal onderwijs. Zeist: ARBO.
- Blank, J. T. L., Boef-van der Meulen, S., Bronneman-Helmers, H. M., Herweijer, L. J., Kuhry, B. & Schreurs, R. A. H. (1990). *School en schaal*. Rijswijk: SCP.
- Boef-van der Meulen, S., Schüssler, I., & Karsten, S. (1993). *Schemergebieden. Het bestuurlijk middenniveau in het primair onderwijs*. Rijswijk: SCP.
- Centraal bureau voor de statistiek (1991). *Statistiek van het basisonderwijs, het speciaal onderwijs en het voortgezet speciaal onderwijs 1990/91. Scholen en leerlingen*. 's-Gravenhage: SDU.
- Centraal bureau voor de statistiek (1992). *Statistiek van het basisonderwijs, het speciaal onderwijs en het voortgezet speciaal onderwijs 1991/92. Scholen en leerlingen*. 's-Gravenhage: SDU.
- Centraal bureau voor de statistiek (1993). *Statistiek van het basisonderwijs, het speciaal onderwijs en het voortgezet speciaal onderwijs 1992/93. Scholen en leerlingen*. 's-Gravenhage: SDU.
- Doornbos, K., & Stevens, L. M. (1987). *De groei van het speciaal onderwijs. Deel A: Analyse van historie en onderzoek*. 's-Gravenhage: SDU.
- Instituut voor sociaal-wetenschappelijk onderzoek (IVA) (1994a). *Terugrapportage WSNS over het schooljaar 1992/93. Een beschrijving van de wijze waarop samenwerkingsverbanden gestalte hebben gegeven aan de zorgplannen. Deel 1: Het taalbeeld*. Tilburg: IVA.
- Instituut voor sociaal-wetenschappelijk onderzoek (IVA) (1994b). *In- en uitstroom van het IOBK*. Tilburg: IVA.
- Keyzer, M. de (1994). *Leerlingstromen tussen het regulier en speciaal onderwijs in de stad Utrecht*. Nijmegen: ITS.
- Meijer, C. J. W. (1988). *Verwijzing gewogen: een studie naar de determinanten van verwijzing naar speciaal onderwijs*. Groningen: RION.
- Meijer, C. J. W. (1995). *Halverwege: Van Startwet naar Streefbeeld*. De Lier: Academisch Boeken Centrum.
- Ministerie van Onderwijs & Wetenschappen (1991). *3 x Akkoord*. 's-Gravenhage: SDU.

Tweede Kamer der Staten-Generaal (1994a). *Weer samen naar school. Memorie van antwoord*, 23486, nr. 5, 14 februari 1994. 's-Gravenhage: SDU.

Tweede Kamer der Staten-Generaal (1994b). *Brief van de Minister van Onderwijs en Wetenschappen aan de Voorzitter van de Tweede Kamer der Staten-Generaal*, 23400 VIII, nr. 63, 22 maart 1994. 's-Gravenhage: SDU.

Wijnstra, J. M. (1987). *Op de grens van basis- en speciaal onderwijs*. Lisse: Swets & Zeitlinger.

Auteur

C.J.W. Meijer Coördinator Evaluatie Weer Samen Naar School.

Adres: Zilver schoon 25, 9801 LM Zuidhorn.

Abstract

Quantitative aspects of the Together to school again-policy

C.J.W. Meijer. *Pedagogische Studiën*, 1996, 73, 404-410.

Recent findings indicate that the Dutch integration policy regarding regular and special education is still waiting for success in terms of the number of children that is integrated in regular settings. The separate special education school system shows an increase of the (absolute) number of children in different schooltypes. Several factors influence the increase of special education. Especially the availability of special education facilities is responsible for this development.

R. H. Hofman en E. B. Vonkeman

Samenvatting

In deze bijdrage wordt ingegaan op de uitkomsten van een onderzoek naar de mate waarin onderwijsmethoden een geschikte conditie vormen voor het realiseren van adaptief onderwijs. In de eerste fase van het onderzoek zijn bijna 50 onderwijsmethoden voor de vakgebieden rekenen, spelling en lezen aan een analyse onderworpen. Verder is een gebruikersonderzoek uitgevoerd onder 91 leerkrachten. De uitkomsten van het onderzoek laten zien dat methoden verschillen in de mate waarin zij een geschikte conditie vormen voor het bereiken van adaptief onderwijs. In de 'adaptieve' methoden zijn basiscomponenten voor adaptief onderwijs goed uitgewerkt, ze bevatten actuele leerinhouden waarmee de kerndoelen voor het basisonderwijs in voldoende mate worden gedekt en zijn goed bruikbaar in combinatieklassen. Deze methoden zijn volgens de leerkrachten bovendien (significant) beter bruikbaar en minder belemmerend voor het realiseren van adaptief onderwijs in hun klas dan de als minst adaptief beoordeelde methoden.

1 Aanleiding en probleemstelling

Uit onderzoek blijkt dat leerkrachten in het basisonderwijs voor vele vak- en vormingsgebieden een onderwijsmethode als leidraad hanteren voor het vormgeven van hun onderwijs (Harskamp & Nijhof, 1987; PPON, 1992). Het reken- en taalonderwijs wordt in de praktijk in grote mate gebaseerd op en beïnvloed door de gebruikte onderwijsmethode (zie o.a. Harskamp, Edelenbos & Suhre, 1993). Of onderwijsmethoden wel voldoende rekening houden met verschillen tussen leerlingen en leerkrachten daadwerkelijk in staat stellen hun onderwijs aan te passen aan de mogelijkheden van leerlingen, is dan ook een belangrijke vraag voor Weer Samen Naar School. Door het GION is

een vergelijkend onderzoek verricht naar de adaptieve kenmerken van onderwijsmethoden waarvan in dit artikel verslag wordt gedaan.

Onderzoek in Nederland laat zien dat de bijdrage van schoolse factoren aan het verklaren van prestaties van leerlingen beperkt is (Cremers, 1991). Factoren die de sociale context van de school representeren verklaren naar schatting meer dan de helft van de variantie in schoolprestaties (Hofman, Hofman & Guldemond, 1995). Het behoeft dan ook geen betoog dat de bijdrage van onderwijsmethoden aan prestaties van leerlingen beperkt is. De geconstateerde smalle marges mogen echter niet leiden tot de conclusie dat onderzoek naar de invloed van onderwijsmethoden van weinig betekenis is. Immers, ook een geringe invloed van (specifieke kenmerken van) onderwijsmethoden op de motivatie, de leerattitude en de leerprestaties van leerlingen kan van verstrekkend belang zijn voor het welbevinden en de schoolloopbaan van een leerling. In verschillende studies zijn de best en slechtst presterende scholen vergeleken en dan blijkt dat leerlingen van de slechtst presterende scholen circa één jaar achter liggen in leerstof of dat er zelfs een verschil van twee leerjaren wordt blootgelegd (Schoor, 1995). Hofman (1993) stelde verschillen vast tussen basisscholen voor het vakgebied rekenen. Het rapportcijfer van een gemiddelde leerling in de 25% slechtst presterende basisscholen bedraagt slechts een vijf, terwijl dat rapportcijfer op de 25% meest effectieve scholen een acht bedraagt (Hofman, 1993). Dus ook geringe schooleffecten leiden tot grote individuele en maatschappelijke gevolgen. Alhoewel de realisatie van adaptief onderwijs en de effecten daarvan op het presteren, de leermotivatie en de zelfbeleving van leerlingen nog niet zo lang onderwerp van onderzoek zijn, kunnen toch al wel enige veelbelovende lijnen worden geschetst. Hieronder wordt allereerst kort ingegaan op het begrip adaptief onderwijs en het empirische materiaal dat over de relatie tussen adaptief onderwijs en de presta-

ties en attitudes van leerlingen beschikbaar is. Vervolgens wordt toegelicht aan welke basiscomponenten een onderwijsmethode dient te voldoen, opdat het materiaal een geschikte conditie vormt voor het realiseren van adaptief onderwijs.

2 Onderwijsmethoden en adaptief onderwijs

Het door de Stuurgroep WSNS geïntroduceerde begrip 'adaptief onderwijs' verwijst naar onderwijs dat is aangepast aan de onderwijsbehoeften van individuele leerlingen met als uiteindelijk doel om het leren voor *alle leerlingen* (risico-leerlingen, modale leerlingen en hoogbegaafde leerlingen) meer effectief te maken (Meijer, Meijnen & Scheerens, 1993). Leerkrachten dienen gevarieerde instructieprincipes te gebruiken om elke leerling de instrumentele vaardigheden te leren beheersen. Adaptief onderwijs dient volgens het procesmanagement WSNS in beginsel groepsgericht te zijn, doch biedt daarbinnen mogelijkheden voor aanpassing aan de specifieke onderwijsbehoeften van individuele leerlingen (Procesmanagement WSNS, 1994; ARBO, 1990). Empirisch onderzoek laat zien dat voor de beheersing van instrumentele basisvakken zoals taal en rekenen, het zogeheten 'directe instructiemodel' uiterst belangrijk is (vgl. Creemers, 1994; Van den Berg & Wolbert, 1993; Veenman, 1992). Uit onderzoek van Gersten en Carnine (1984) naar de effectiviteit van het directe instructie-model voor onderwijs aan leerlingen uit achterstandssituaties blijkt dat na afloop van een vierjarige interventie de rekenvaardigheid van de achterstandsleerlingen vergelijkbaar is met die van leerlingen uit de niet-achterstandssituatie, het zelfbeeld versterkt is en het gevoel voor verantwoordelijkheid voor de eigen leerprestaties toegenomen is. Houtveen en Booij (1994) concluderen eveneens dat in diverse onderzoeken in het angelsaksische taalgebied evidentie is gevonden voor de effectiviteit van vormen van adaptief onderwijs bij het realiseren van verbetering van leerlingresultaten (Gersten & Carnine, 1984; Corno & Snow, 1986; Walberg & Wang, 1986; Slavin, 1994).

Basiscomponenten van adaptief onderwijs

De relatie tussen adaptief onderwijs en onderwijsmethoden en de effecten daarvan op leerprestaties van leerlingen is een onderzoeks-terrein dat pas recent aandacht krijgt. Het schaarse beschikbare empirische materiaal laat zien dat bij beoordeling van onderwijsmethoden als conditie voor adaptief onderwijs de instrumentering van een zestal basiscomponenten centraal dienen te staan (zie: Van der Leij, 1993; Walberg & Wang, 1986; Corno & Snow, 1986; Houtveen, Booij & De Jong, 1995):

- instrumenten voor signalering van achterstanden en/of leerproblemen en het opsporen van risicoleerlingen;
- instrumenten voor het diagnostiseren en het verder lokaliseren van specifieke leerproblemen (aard en uitval op leerstofonderdelen);
- instrumenten voor differentiatie in leerwegen;
- instrumenten voor variatie in instructiewijzen om de instructie af te stemmen op verschillende behoeften van de leerlingen;
- instrumenten voor toetsing en evaluatie;
- instrumenten voor remediëring: aanwijzingen voor gebruik van remedieel materiaal om de achterstanden van zwakke leerlingen te verminderen.

Het gaat hierbij overigens om aspecten die elkaar overlappen. De beschikbare empirische studies over onderwijsmethoden zijn geanalyseerd met het oog op voornoemde relevante basiscomponenten voor het bereiken van adaptief onderwijs. In vogelvucht worden de belangrijkste implicaties daarvan voor onderwijsmethoden gepresenteerd.

Wat *signaleren* en *diagnostiseren* betreft dienen in onderwijsmethoden signalerings- of instaptoetsen aanwezig te zijn die worden afgenomen voordat leerlingen met een nieuw onderdeel beginnen. Bovendien moet daarbij expliciet worden aangegeven hoe met de informatie uit de toetsafname een gedifferentieerd onderwijsaanbod kan worden aangeboden. Formele technieken van signalering hebben daarbij de voorkeur boven meer informele technieken (vgl. Janssens, 1986; Pijl & Wolfgram, 1990). Hierop aansluitend is het vanzelfsprekend dat het onderwijs op de geconstateerde verschillen tussen leerlingen wordt aangepast: *differentiatie in leerwegen*. Onder-

wijsmethoden dienen voldoende differentiatie-mogelijkheden te bevatten voor alle leerlingen, opdat zij voldoende tijd en mogelijkheden hebben om de gestelde taak te leren (Reezigt, Van Dijk & Bosveld, 1986; Wolbert et al., 1986). Over het effect van het toepassen van verschillende leerstofdifferentiatie-modellen zijn geen eenduidige conclusies te trekken (Reezigt 1993, Guldemon, 1994). Het werken met niveaudifferentiatie is echter niet in het voordeel van risicoleerlingen, omdat het niet bevorderlijk is voor de leermotivatatie van deze kinderen en niet leidt tot een meer positieve inschatting van hun prestaties door de leerkracht dan bij heterogeen groeperen het geval is. De kern van adaptief onderwijs vormt vanzelfsprekend de instructie, meer specifiek de *variatie in instructiewijzen* die de leerkracht hanteert. In recente onderzoeken zijn modellen voor effectieve instructie beproefd. De indirecte instructiestijl (ook interactieve instructie genoemd), wordt vooral effectief geacht voor het leren toepassen van nieuw opgedane kennis en vaardigheden (vgl. Van den Berg & Wolbert, 1993; Van der Leij, 1993). De directe instructie lijkt vooral effectief te zijn voor de instrumentele basisvakken en met name voor zwakkere leerlingen. Over het algemeen is de conclusie uit onderzoek (zie Veenman, 1992) dat risicoleerlingen door middel van stapsgewijze instructie en feedback aan de hand van oplossingschema's tot betere leerprestaties komen dan alleen via verbale uitleg. Onderzoek naar *toetsing en evaluatie van vorderingen* laat zien dat een in de methode opgenomen toetsstructuur (met regelmatige toetsafname), een belangrijk middel vormt voor de leerkracht om adaptief onderwijs te verzorgen (Reezigt, 1993; Kool, 1990). Reezigt (1993) geeft aan dat idealiter de toetsen in de stof geïntegreerd zouden moeten zijn en dat correctief materiaal en verrijkingsstof gekoppeld moeten zijn aan de toetsresultaten van leerlingen. Ook zou het materiaal zich meer moeten lenen voor zelfstandig gebruik door leerlingen (Reezigt, 1993; Janssens, 1986; Klaasman, 1989; Slavenburg & Peters, 1989). Daarnaast concludeert Brandsma (1993) dat het regelmatig afnemen van gestandaardiseerde schoolvorderingentoetsen en het centraal administreren van prestatiegegevens belangrijke instrumenten kunnen zijn om de prestaties van leerlingen positief te beïnvloe-

den. In het kader van de evaluatie van effecten van *remediële programma's* voor kinderen met achterstanden in spellingvaardigheid komt Assink (1992) tot de conclusie dat een strak opgebouwde leergang gebaseerd op regels voor het spellen die zijn voorzien van schema's waarmee instructie en feedback wordt gegeven, tot betere resultaten leidt dan de traditionele leergangen voor spelling. Willemsen (1994) deed onderzoek naar het effect van een leergang remediële rekenen. Ook hier blijkt de regelgeleide en visueel ondersteunde leergang die voorzien is van directe feedback, tot betere resultaten te leiden dan andere remediële leergangen. Willemsen wijst er overigens op dat de remediële methodiek zo min mogelijk dient af te wijken van de methodiek die in de gewone rekenmethode wordt gevolgd.

3 Ontwerp van het onderzoek

Bij het verzorgen van adaptief onderwijs binnen de klas is het uitgangspunt dat de leerling uiteindelijk de instrumentele basisvaardigheden beheerst. In dit onderzoek ligt het accent om die reden op de vakgebieden lezen, spellen en rekenen in de onderbouw. Het gaat hierbij om de vakgebieden en de leerjaren waar het vaakst leerproblemen gesignaleerd worden op grond waarvan doorverwijzing naar het speciaal onderwijs plaatsvindt (Meijer, 1988). Het onderzoek beslaat drie fasen waarin telkens de vraag centraal staat in welke mate gangbare en nieuwe onderwijsmethoden binnen deze vakgebieden het mogelijk maken adaptief onderwijs te realiseren. De eerste fase betreft een screening door onderzoekers gespecialiseerd in methodebeoordelingen van een groep van bijna 50 onderwijsmethoden voor de vakgebieden rekenen ($n = 20$), spelling ($n = 15$) en lezen ($n = 12$). De tweede onderzoeksfase betreft de beoordeling door deskundigen van een selectie van zes methoden per vakgebied die als beste uit de screening kwamen.¹ Als derde onderzoeksfase is een gebruikersonderzoek uitgevoerd waarin het gebruik door leerkrachten ($n = 91$) van een methode die als het meest geschikt en een methode die als het minst geschikt is beoordeeld voor het realiseren van adaptief onderwijs in de klas (voor drie vakgebieden).

Bij de constructie van instrumenten stonden twee concepten centraal, namelijk (a) de al eerder besproken zes basiscomponenten van adaptief onderwijs en (b) de eigentijdsheid van de onderwijsmethode.² Dit laatste begrip verwijst naar het oordeel van de Commissie Evaluatie Basisonderwijs (1994) die stelt dat pas sprake is van een goed onderwijsaanbod als de methode maatschappelijk relevant en voldoende eigentijds is en de kerndoelen voor het basisonderwijs voldoende dekt (zie Harskamp & Suhre, 1993).

Operationalisering van adaptief onderwijs

Bij de operationalisatie van het construct adaptief onderwijs gaat het om zes groepen van items. Hieronder worden de specifieke onderdelen van de geconstrueerde schalen (operationalisaties) kort geëxpliciteerd. Voor meer informatie over de schalen wordt verwezen naar het wetenschappelijk rapport van de studie (Hofman & Vonkeman, 1995a).

Bij *signalering* is nagegaan in hoeverre in de methode signalerings- en observatie-instrumenten en instaptoetsen aanwezig zijn. Verder is onderzocht of er concrete aanwijzingen zijn voor een frequente afname van de toetsen/instrumenten alsmede voor evaluatie van leerresultaten tijdens lesobservaties en analyse van het werk van de leerlingen. Verder is gelet op criteria om prestaties tegen af te zetten, alsmede aanwijzingen voor het inschatten van niveauverschillen en het bijhouden van een leerlingvolgsysteem. Bij *diagnose* wordt een onderscheid gemaakt tussen mondeling diagnostisch onderzoek en het schriftelijk afnemen en analyseren van toetsen. Voor het mondelinge onderzoek kunnen aanwijzingen worden gegeven voor het opsporen van foute oplossingsstrategieën (bijvoorbeeld in een leergesprek). Voor het schriftelijk onderzoek kunnen aanwijzingen worden gegeven voor foutenanalyse en vaststellen van foute strategieën bij leerlingen. Bij *differentiatie in leerwegen* is gekeken naar verschillende leerstoforganisatiemodellen. Centraal staat het flexibel en heterogeen groeperen met als uitgangspunt de klas als geheel. Gelet is op aanwijzingen voor variatie in groeperingswijzen (ook in de tijd). Bij *variatie in instructiewijzen* worden flexibele instructiewijzen positief beoordeeld. Er is nagegaan of concrete aanwij-

zingen in de methoden zijn opgenomen voor toepassing van instructie op grond van uitkomsten uit de signalering of diagnose en of er aanwijzingen zijn voor directe instructie voor zwakke leerlingen én interactieve instructie voor betere leerlingen. Bovendien is nagegaan of er sprake is een specifieke opbouw van de instructie en of de methode expliciet aandacht schenkt aan positieve feedback naar leerlingen. Bij *toetsing en evaluatie* is gelet op regelmaat in het toetsen en het soort toetsen: toetsen met beheersingsnorm, speciale toetsingscriteria voor de minimumleerstof, voldoende toetswerkbladen of dictees en zelfcontroleerende opdrachten/werkbladen. Verder is nagegaan of in de methode een registratiesysteem aanwezig is waarin toets- en andere gegevens over de leerlingen met regelmaat kunnen worden vastgelegd en of er leerlinglijsten worden gegeven voor de overdracht van gegevens aan andere leerkrachten.

Bij *remediëring* staat de verwijzing op grond van signalering/diagnose naar leerstof voor oefening en herhaling door de zwakke leerling met specifieke problemen centraal. Gelet is op remediëring, zoveel mogelijk binnen klasseverband, en met aanwijzingen voor nabespreking van het remediële werk. Verder is nagegaan of in de onderwijsmethode aanwijzingen aanwezig zijn voor het maken van een handlingsplan alsmede voor de klasseorganisatie bij remediële hulp in de klas.

Operationalisering van eigentijdsheid

Er is een inschatting gevraagd omtrent de actualiteit van de onderwijsmethode, de hoeveelheid leerstof in de methode voor de basisvaardigheden (per vakgebied) en de bijdrage die de onderwijsmethode naar het oordeel van de deskundige levert aan het op termijn (eind groep 8) realiseren van de kerndoelen voor het basisonderwijs. De items van deze drie schalen zijn voor een deel ontleend aan een deelstudie die is uitgevoerd in opdracht van de Commissie Evaluatie Basisonderwijs (Harskamp & Suhre, 1993). Daarnaast is een apart item omtrent het gebruik van de onderwijsmethode in combinatieklassen opgenomen.

Psychometrische gegevens

Op basis van de uitgevoerde betrouwbaarheidsonderzoeken mag worden geconcludeerd

dat zowel bij de screening, het deskundigen-onderzoek als het gebruikersonderzoek sprake is van een voldoende tot hoge mate van interne consistentie bij de meting van de basiscomponenten van adaptief onderwijs ($\alpha = >.90$). Alleen voor het vakgebied rekenen zien we op onderdelen lagere betrouwbaarheden (voor diagnose, remediëring en toetsing). Ook de schalen betreffende de eigentijdsheid van de onderwijsmethode en de aansluiting bij kern-doelen voor het basisonderwijs zijn betrouwbaar.

4 Resultaten van het methoden-onderzoek

Screening van reken-, spelling- en lees-methoden

Drie soorten van onderwijsmethoden en materialen zijn gescreend, namelijk basismethoden,

remediële methoden en aanvullende materialen of oefenstof. De laatste twee zijn vooral gescreend met het oog op het flexibiliseren van gangbare onderwijsmethoden. Uiteindelijk zijn voor rekenen 16 methoden (acht basismethoden en acht remediële methoden) en vier pakketten met aanvullende materialen of oefenstof geselecteerd. Voor spelling betreft het in totaal 14 methoden: zes basismethoden voor taalonderwijs, zes specifieke spellingleergangen en één pakket aanvullende oefenstof. Voor lezen betreft de screening tien basismethoden – drie voor aanvankelijk en twee voor voortgezet technisch lezen en vijf voor begrijpend lezen – en twee remediële methoden. Elke methode is geanalyseerd op de mate waarin de basiscomponenten voor adaptief onderwijs zijn uitgewerkt. Hiervan zijn uitgebreide beschrijvingen gemaakt die zijn opgenomen in het zogenaamde *Bronnenboek Adaptief Onderwijs* (Hofman & Vonkeman, 1995b). Per basiscom-

Tabel 1
Screening rekenmethoden op basiscomponenten adaptief onderwijs

screening methoden rekenonderwijs	sign	dia	dif	ins	re	toe	totaal
<i>Basismethode</i>							
Rekenwerk	++	++	□	++	□	+	(zeer) goed
Pluspunt	++	□	++	++	+	+	(zeer) goed
De wereld in getallen ¹	+	□	+	++	□	□	goed
Rekenen & Wiskunde	+	-	□	+	□	□	voldoende
Operatoir rekenen gr.3 ¹	+	+	□	+	--	+	goed
Idem groep 4/5 ¹	□	□	--	□	--	□	matig
Getal in beeld	--	--	--	--	--	--	onvold.
Taltaal	--	--	--	--	--	--	onvold.
Naar zelfstandig rekenen	--	--	--	--	--	--	onvold.
<i>Remediële methode</i>							
Kwantiwijzer ¹	++	++	+	++	++	+	(zeer) goed
Gouds Rekenpakket	+	++	++	+	□	□	(zeer) goed
Diagn. Rekenonderzoek	nvt	++	+	++	+	+	(zeer) goed
RekenRadar	++	++	+	++	□	□	(zeer) goed
Rekenspoor	+	+	□	+	□	--	goed
Zo reken ik ook!	+	--	□	+	□	--	vold./goed
Remelka	-	--	+	++	nvt	--	voldoende
Ik weet hoe ik reken	--	--	--	--	--	--	onvold.

1. nieuwste versie scoreverdeling: ++ = zeer goed
+ = goed
□ = matig
- = voldoende
-- = onvoldoende

ponent van adaptief onderwijs is een oordeel gegeven: van onvoldoende, matig, voldoende, goed tot zeer goed uitgewerkt in de methode. Op basis van deze scores is vervolgens een eindoordeel over de methode berekend. De resultaten van de screening worden hieronder per vakgebied afzonderlijk besproken. In Tabel 1 worden ter illustratie de gegevens voor het vakgebied rekenen gepresenteerd. Voor overzichten voor spelling en lezen wordt verwezen naar Hofman & Vonkeman, 1995a).

Van de acht gescreende basismethoden voor **rekenonderwijs** worden *Pluspunt*, *Rekenwerk* en *De Wereld in Getallen* op alle punten (vooral signalering en instructie) goed tot zeer goed uitgewerkt beschouwd voor het realiseren van adaptief onderwijs. De rekenmethode *Operator Rekenen* is niet eenduidig uitgewerkt: het materiaal voor groep 3 is goed uitgewerkt, voor groep 4/5 matig. Drie methoden zijn als onvoldoende adaptief gescreend: alle zes basiscomponenten zijn onvoldoende uitgewerkt. Het betreft oudere rekenmethoden, die op dit moment veelal worden vervangen door modernere rekenmethoden. Vijf van de acht remediële methoden krijgen het predikaat goed tot zeer goed uitgewerkt voor het realise-

ren van adaptief onderwijs, terwijl slechts één remediële methode als onvoldoende wordt gescreend. Vooral de (recent uitgegeven) remediële rekenmethoden scoren zeer hoog op bijna alle basiscomponenten.

Van de **spellingleergangen** die zijn opgenomen in een basismethode voor het taalonderwijs wordt alleen *Taal Actief (nieuwe versie)* als zeer goed beoordeeld. Verder zijn drie basismethoden voldoende uitgewerkt, namelijk *Taal Actief (oude versie)*, *Taalkabaal* en *Taal-journaal*. Drie basismethoden zijn slechts matig uitgewerkt op de basiscomponenten: *Taal Totaal*, *Spellingland* en *Jouw Taal, Mijn Taal*. Van de speciale spellingleergangen zijn het *Spellingspakket* en *Spelling in de Lift* zeer goed uitgewerkt, terwijl *Woordbouw* en *Taal is niet zo moeilijk* goed zijn uitgewerkt op de componenten van adaptief onderwijs. *Spelkwartier* is een remediële spellingleergang die het predikaat matig uitgewerkt krijgt.

Van de 12 **leesmethoden** die zijn gescreend worden er slechts twee beschouwd als methoden waarin de basiscomponenten voor adaptief onderwijs zeer goed zijn uitgewerkt: *Leeslijn/De Leesweg* en *Leessleutel* (beide recent uitgegeven methoden voor aanvankelijk technisch

Tabel 2

Adaptief onderwijs: best uitgewerkte methoden voor drie vakgebieden

Basismethoden	oordeel screening	actualiteit ¹	kerndoelen ²	combinatiekl. ³
rekenen				
Pluspunt	zg	3.7	2.5	4.0
De Wereld in Getallen ⁶	g	3.3	2.4	2.7
Rekenwerk	zg	3.0	2.6	1.4
Rekenen & Wiskunde	v	3.0	2.3	2.3
spelling				
Taal Act. Woordspel (nieuw) ⁴	zg			
Taal Act. Woordspel (oud)	v	2.5	2.0	2.8
Woordbouw	g	2.3	2.3	3.3
lezen				
Leeslijn/De Leesweg ⁵	zg	3.1	2.5	3.2
Leessleutel ⁶	zg	2.8	2.1	2.8
Veilig Leren Lezen (m-r-v) ⁶ differentiatiemateriaal	v/g	2.9	1.8	2.3
Leeswerk (voortg. techn. I.)	g	3.2	2.2	2.3
Lees Je Wijzer (begr. lez.)	m	3.0	2.7	3.0

zg = zeer goed, g = goed, v = voldoende, m = matig, o = onvoldoende;

1 scores lopen van 1=voor vrijwel geen enkele les/zelden tot 4=voor vrijwel alle lessen/in sterke mate

2 scores lopen van 1 = geringe bijdrage tot 3 = grote bijdrage

3 scores lopen van 1 = nee, nauwelijks tot 4 = in sterke mate uitvoerbaar in combinatieklassen

4 van deze methode zijn op dit moment alleen de spellingmaterialen voor groep 4 en 5 beschikbaar

5 van deze methode ontbreekt nog een beperkte hoeveelheid materialen

6 bij deze methoden is software verkrijgbaar

lezen). De veel gebruikte methode *Veilig Leren Lezen* wordt als voldoende tot goed beoordeeld voor het realiseren van adaptief onderwijs. Van de methoden voor voortgezet technisch lezen wordt *Leeswerk* goed uitgewerkt geacht. Bij de methoden voor begrijpend lezen krijgt *Begrijpend lezen* het predikaat goed uitgewerkt voor het realiseren van adaptief onderwijs. Van de twee beoordeelde remediële leesmethoden krijgt *Leesspoor* het predikaat goed uitgewerkt op de basiscomponenten van adaptief onderwijs, terwijl *Speciale leesbegeleiding* als voldoende uitgewerkt wordt beoordeeld.

Uit de hiervoor besproken vakgebieden komt naar voren dat onderwijsmethoden verschillen in de mate waarin daarin kenmerken zijn opgenomen die adaptief onderwijs in de klas mogelijk maken. In Tabel 2 zijn de bevindingen uit de screening van de best beoordeelde methoden, voor de drie vakgebieden, naast elkaar gezet.

Duidelijk is dat al de genoemde methoden een geschikte voorwaarde vormen voor het realiseren van adaptief onderwijs. Niet elke school is op dit moment in staat om een nieuwe methode aan te schaffen.³ De financiën ontbreken of er is (relatief) kort geleden een nieuwe methode aangeschaft. De onderzoekers bevelen daarom bovendien aanvullende methoden of deelleergangen per vakgebied aan, die mogelijkheden bieden tot flexibilisering van minder adaptieve methoden (zie Hofman & Voncken, 1995a, 1995b).

Methodegebruik en leerkrachtgedrag

Aan het gebruikersonderzoek hebben 91 leerkrachten van groep 3, 4 en/of 5 van 50 scholen hun medewerking verleend. Een representativiteitsonderzoek (aan de hand van landelijke referentiegegevens laat zien dat de onderzoeksgroep nauwelijks afwijkt van de populatie van Nederlandse basisscholen.⁴

Voor het gebruikersonderzoek zijn zes methoden geselecteerd, namelijk de best en slechtst gescreende methode per vakgebied. Voor rekenen zijn dit respectievelijk *Pluspunt* (als meest adaptief gescreend) en *Naar Zelfstandig Rekenen* (minst adaptief). Voor spelling betreft het *Taal Actief* (oude versie; de nieuwe versie werd beter beoordeeld, maar is nog te kort op de markt) en *Jouw Taal, Mijn Taal* (minst adaptief). Van de leesmethoden

zijn de methoden *Leeslijn/De Leesweg* en *Lees Je Wijzer* gekozen, respectievelijk als meest en minst adaptieve methode. Overigens gaat het hier om een gebruik in groep 3/4 en/of 5 van een methode voor aanvankelijk technisch lezen versus een methode voor begrijpend lezen. Het bleek onmogelijk een best en slecht gescreende leesmethode voor óf technisch óf begrijpend lezen te vergelijken, vandaar de selectie van dit tweetal.

De uitkomsten van het gebruikersonderzoek laten zien dat er voor **spelling en lezen** in het algemeen nauwelijks verschillen zijn tussen de leerkrachten in de toepassing van de basiscomponenten. Slechts één keer wordt een significant verschil geconstateerd: leerkrachten die de meest adaptieve leesmethode gebruiken, blijken vaker de instructie en de leertijd te variëren, al naar gelang de behoeften van de leerlingen, dan leerkrachten die de minst adaptieve leesmethode gebruiken.

Voor het vakgebied **rekenen** is het beeld evenwel anders. De gebruikers van de minst adaptieve methode *Naar Zelfstandig Rekenen* scoren op toetsing en evaluatie significant lager, maar op diagnostiek juist significant hoger dan de gebruikers van de meest adaptieve methode *Pluspunt*. Bij de interpretatie hiervan is echter voorzichtigheid geboden, omdat slechts 41% van de gebruikers van *Naar Zelfstandig rekenen* de methode volledig gebruikt, versus 100% van de gebruikers van *Pluspunt*. Bovendien valt op dat de gebruikers van *Naar Zelfstandig Rekenen* voor significant kleinere klassen staan dan *Pluspunt* gebruikers.

Uit het oogpunt van het gebruikersperspectief is de leerkrachten bovendien gevraagd aan te geven hoe bruikbaar enerzijds en hoe belemmerend anderzijds de door hen gebruikte methode is voor het realiseren van adaptief onderwijs in de klas. De achterliggende gedachte bij deze vraagstelling is dat het niet altijd zo hoeft te zijn dat een als weinig bruikbaar beoordeelde methode, ook altijd belemmerend werkt voor het bereiken van adaptief onderwijs. Per methoden is de gemiddelde score in Tabel 3 weer gegeven.

Er worden (significante) verschillen tussen alle meest en minst adaptieve methoden vastgesteld. De gebruikers van de meest adaptieve methoden beoordelen hun methoden als beter bruikbaar én minder belemmerend voor het

Tabel 3*Bruikbaarheid/belemmering van methoden voor bereiken van adaptief onderwijs*

	bruik- baarheid	belemme- ring
<i>rekenmethoden</i>		
Pluspunt (meest adaptief)	2.5*	3.7*
Naar Zelfst. Rek. (minst)	1.4	2.3
<i>spellingmethoden</i>		
Taal Actief (meest)	2.0*	3.8*
Jouw Taal, Mijn Taal (minst)	1.3	2.7
<i>leesmethoden</i>		
Leeslijn/De Leesweg (meest)	2.5*	3.8*
Lees Je Wijzer (minst)	1.6	2.6

* het verschil tussen beide gemiddelde scores is significant
scoreverdeling: 1 = onvoldoende bruikbaar/sterk belemmerend tot 4 = zeer goed bruikbaar/niet belemmerend

realiseren van adaptief onderwijs, dan de gebruikers van de minst adaptieve methoden. Overigens valt op dat een relatief minder bruikbare methode niet altijd even belemmerend behoeft te zijn voor het voor het bereiken van adaptief onderwijs. De rekenmethode Naar Zelfstandig Rekenen is qua bruikbaarheid vergelijkbaar met de andere niet adaptieve methoden, maar wordt wel als meer belemmerend beoordeeld. Gegeven de eerder genoemde bevindingen uit het gebruikersonderzoek betreffende de twee rekenmethoden lijkt het er hier eveneens op dat Naar Zelfstandig Rekenen een leerkracht dwingt om materialen uit andere methode toe te voegen dan wel zelf bij te maken, om toch goed rekenonderwijs te geven. Overigens stelden onderzoekers naar implementatie en gebruik van onderwijsmethoden reeds in 1982 vast dat gedetailleerd uitgewerkte onderwijsmethoden in hogere mate volledig worden geïmplementeerd, dan weinig gedetailleerde onderwijsmethoden (Harskamp & Hofman, 1982).

5 Aanbevelingen en discussie

Het onderzoek naar de onderwijsmethode als conditie voor adaptief onderwijs is slechts een eerste stap om meer zicht te krijgen op geschikte onderwijsmethoden voor het realiseren van adaptief onderwijs. Overigens kent dit onderzoek ook enige beperkingen die hier kort worden aangestipt.

Het onderzoek beperkt zich tot een niet-

vakspecifieke beoordeling van adaptiviteit van de methode. Er is bijvoorbeeld wel gevraagd of er minimumdoelen voor alle leerlingen in de methode zijn gegeven, maar niet welke minimumdoelen dat dan zijn. Overigens wordt daaraan wel aandacht besteed in de uitgebreide afzonderlijke beschrijvingen van elke beoordeelde onderwijsmethode in het *Bronnenboek Adaptief Onderwijs* (Hofman & Vonkeman, 1995b).

Een andere belangrijke beperking is dat het gebruikersonderzoek een schriftelijke vragenlijst betrof en geen observatie-onderzoek onder leerkrachten. Sociale wenselijkheid kan de resultaten beïnvloeden en het zou kunnen zijn dat dat bij leerkrachten die relatief weinig kenmerken van adaptief onderwijsgedrag vertonen, sterker het geval is. Bovendien is nog onduidelijk wat de effecten zijn van het niet volledig gebruiken van een methode en het gebruik van aanvullende materialen. Onderzoek laat zien dat aanvullend gebruik meer voorkomt bij oudere methoden dan bij de recent ontwikkelde methoden (Harskamp & Suhre, 1993; PPON, 1992). In dit gebruikersonderzoek is bovendien niet de onderwijsleersituatie van bijvoorbeeld het rekenonderwijs onderzocht, maar alleen de wijze van gebruik van de rekenmethode. Overigens is de bevinding dat de als adaptief beoordeelde methoden in het algemeen volledig worden gebruikt en de als niet-adaptief beoordeelde slechts gedeeltelijk, interessant vanuit een gebruikersperspectief. Onderzoek laat zien dat het effect van een methode staat of valt met het gebruik ervan door de leerkracht. Een methode waarin nauwelijks instrumenten zijn opgenomen om basiscomponenten van adaptief onderwijs te realiseren zou daarmee een belemmering vormen voor het realiseren van adaptief onderwijs. Wil een leerkracht die een als minder adaptief beoordeelde methode gebruikt toch effectief onderwijs bereiken, dan zal hij of zij zich extra moeite moeten getroosten om via aanvullende (bijv. zelfvervaardigde) materialen en instrumenten het onderwijsproces te optimaliseren. De verwachting dat een methode waarin nauwelijks instrumenten zijn opgenomen om basiscomponenten van adaptief onderwijs te realiseren, een belemmering zou vormen voor het realiseren van adaptief onderwijs door de leerkracht, wordt in ieder geval bevestigd door de resultaten van ons gebruikerson-

derzoek. Omdat de taakbelasting van leerkrachten in het huidige onderwijs niet gering is, lijkt het verstandig methoden voor adaptief onderwijs aan te bevelen waarin zoveel mogelijk materialen, instrumenten en concrete aanwijzingen zijn opgenomen die adaptief leerkrachtgedrag kunnen stimuleren.

Overigens beschikken we (nog) niet over empirische gegevens om de relatieve bijdrage van onderwijsmethoden aan het realiseren van adaptief onderwijs in kwantitatieve termen te kunnen schatten. Daarvoor is onderzoek naar de bijdrage van interacties tussen leerkrachtgedrag en kenmerken van onderwijsmethoden noodzakelijk. Met name het vaker vastgestelde verschil in (volledig) gebruik van onderwijsmethoden is daarbij interessant (Creemers, 1991; Nijhof et al., 1993). In Nederland gebruiken scholen een grote variëteit aan onderwijsmethoden. Daarom is het zinvol om vergelijkingen te maken tussen de methoden onderling en aldus het relatieve effect ervan, alsmede interacties met leerkracht- en leerlingkenmerken op het realiseren van adaptief onderwijs en indirect op de prestaties en het welbevinden van leerlingen vast te stellen.

Noten

- 1 Vanwege de beperkte ruimte wordt dit deelonderzoek verder niet besproken. Overigens blijven de oordelen van de deskundigen in hoge mate te stroken met die van de screening: zelfde rangorde, maar de oordelen zijn wat strenger.
- 2 Het Procesmanagement Weer Samen Naar School heeft op basis van haar definitie van adaptief onderwijs negen kwaliteitscriteria – de zogenaamde 'WSNS-standaards' – voor adaptief onderwijs op groepsniveau onderscheiden (zie Procesmanagement WSNS, 1994). Ook daaraan is aandacht geschonken in dit onderzoek (zie Hofman & Vonkeman, 1995a).
- 3 In het Bronnenboek Adaptief Onderwijs worden in de tabel ook de aanschaffkosten weergegeven van eenmalige aanschaf van een set van alle beschikbare materialen (inclusief een set werk/toets-boeken indien van toepassing, de omvang van deze sets is veelal vijf exemplaren).
- 4 Er is sprake van lichte afwijkingen: de steekproef bevat wat meer kleine scholen, meer openbare scholen, vaker combinatieklassen.

Literatuur

- ARBO (1990). *Opmaat tot samenspel. Advies over de binding tussen regulier en speciaal onderwijs*. Zeist: ARBO.
- Assink, E. M. H. (1992). Leren spellen. In L. Verhoeven (red.), *Handboek lees- en schrijfdidactiek* (pp. 184-195). Lisse: Swets & Zeitlinger.
- Berg, G. van den, & Wolbert, R. G. M. (1993). *Vormgeving en effecten van het onderwijsaanbod in het basisonderwijs*. Groningen: RION/RUG.
- Brandsma, H. P. (1993). *Basisschoolkenmerken en de kwaliteit van het onderwijs*. Academisch proefschrift. Groningen: RION/RUG.
- Commissie Evaluatie Basisonderwijs (1994). *Zicht op kwaliteit. Eindrapport*. De Meern: Inspectie van het onderwijs.
- Corno, L., & Snow, R. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Ed.), *Third handbook of research on teaching*. Washington: American Educational Research Association.
- Creemers, B. P. M. (1991). *Effectieve instructie. Een empirische bijdrage aan de verbetering van het onderwijs in de klas*. Den Haag: SVO. Balansreeks.
- Creemers, B. P. M. (red.) (1994). *Deregulering en de kwaliteit van het onderwijs*. Groningen: GION/RUG.
- Gersten, R., & Carnine, D. (1984). Direct instruction in mathematics: a longitudinal evaluation of low-income elementary school students. *Elementary School Journal*, 84, 395-406.
- Guldemond, H. (1994). *Van de kikker en de vijver: groepseffecten op individuele leerprestaties*. Academisch proefschrift. Groningen: GION/RUG.
- Harskamp, E. G., & Hofman, R. H. (1982). *Evaluatie van leermethoden in het basisonderwijs*. Haren: RION.
- Harskamp, E. G., & Nijhof, W. J. (1987). *Curriculum: Ontwerp, implementatie en evaluatie*. Lisse: Swets & Zeitlinger.
- Harskamp, E. G., & Suhre, C. J. M. (1993). *Beoordeling van onderwijsmethoden in het basisonderwijs*. Groningen: RION/RUG.
- Harskamp, E. G., Edelenbos, P., & Suhre, C. J. M. (1993). *Feitelijk aanbod voor de vakken Nederlandse Taal, Rekenen en Wiskunde en Aardrijkskunde op de basisschool (concept eindrapport)*. Groningen: RION/RUG.
- Hofman, R. H. (1993). *Effectief schoolbestuur. Een studie naar de bijdrage van schoolbesturen aan de effectiviteit van basisscholen*. Academisch proefschrift. Groningen: RION/RUG.

- Hofman, R. H., Hofman, W. H. A., & Guldemond, H. (1995). *Governance, schools and classrooms and their impact on effectiveness in private and public schools*. Paper presented at ECER, Bath, England. Submitted to American Educational Research Journal.
- Hofman, R. H., & Vonkeman, E. B. (1995a). *Conditioes voor adaptief onderwijs: de onderwijsmethode*. Wetenschappelijk onderzoeksrapport. Groningen: GION/RUG.
- Hofman, R. H., & Vonkeman, E. B. (1995b). *Conditioes voor adaptief onderwijs: de onderwijsmethode. Bronnenboek adaptief onderwijs. Onderwijsmethoden voor rekenen, spelling en lezen op de basisschool*. Groningen: GION/RUG.
- Houtveen, A. A. M., & Booij, N. (1994). *Het meten van integrale leerlingzorg: adaptief onderwijs en schoolontwikkeling*. Utrecht: ISOR/Afdeling Onderwijsonderzoek, Universiteit Utrecht.
- Houtveen, A. M. M., Booij, N., & Jong, R. de (1995). *Effectieve begeleiding en leerkrachtgedrag*. Utrecht: ISOR/Afdeling Onderwijsonderzoek Universiteit Utrecht.
- Janssens, F. J. G. (1986). *De evaluatiepraktijken van leerkrachten. Een beschrijvend onderzoek naar het evalueren tijdens het rekenen in het primair onderwijs*. Arnhem: CITO.
- Klaasman, R. R. P. (1989). *Programma-implementatie en schoolprestaties*. Project Onderwijs en Sociaal milieu. Rotterdam: SAD.
- Kool, E. (1990). *Werken volgens plan in het speciaal onderwijs*. Academisch proefschrift. Amsterdam: Vrije Universiteit.
- Leij, A. van der (1993). Zorgverbreding: stand van zaken en perspectief. In A. van der Leij & E. J. Kappers (1993). *Zorgverbreding: bijdragen uit het speciaal onderwijs aan basisonderwijs* (pp. 296-321, 5e herziene druk). Nijkerk: Intro.
- Meijer, C. J. W. (1988). *Verwijzing gewogen: een studie naar de determinanten van verwijzing naar speciaal onderwijs*. Groningen: RION.
- Meijer, C. J. W. (1990). Onderzoek naar zorgverbreding: thema's en methoden. In M. K. van der Heijden & J. Rispen (red.), *Zorgverbreding en Afgrenzing* (pp. 87-96). Amsterdam/Lisse: Swets & Zeitlinger.
- Meijer, C., Meijnen, W., & Scheerens, J. (1993). *Over wegen en schatten en sturen*. Analytische beleidsevaluatie Weer Samen naar School. De Lier: Academisch Boeken Centrum.
- Nijhof, W. J., Franssen, H. A. M., Hoebe, W. Th. J. G., & Wolbert, R. G. M. (red.) (1993). *Handboek curriculum. Modellen, theorieën, technologieën*. Lisse: Swets & Zeitlinger.
- PPON (J. Sijstra, red.) (1992). *Balans van het taalonderwijs halverwege de basisschool. Uitkomsten van de eerste taalpeiling medio basisonderwijs*. Arnhem: Cito. PPON-reeks, nr. 3.
- Procesmanagement Weer Samen Naar School (1994). *Basisdocument WSNS-Standaards. Theoretische achtergrond en referentiekader*. 's-Hertogenbosch: Procesmanagement WSNS.
- Pijl, S. J., & Wolfgram, H. P. (1990). Evalueren binnen nauwe grenzen. In M. K. van der Heijden & J. Rispen (red.), *Zorgverbreding en afgrenzing* (pp. 59-68). Lisse: Swets & Zeitlinger.
- Reezigt, G. J. (1993). *Effecten van differentiatie op de basisschool*. Academisch proefschrift. Groningen: RION/RUG.
- Reezigt, G. J., Dijk, M. H. van, & Bosveld, J. J. F. (1986). *Differentiatie op de basisschool*. Den Haag: SVO/RION.
- Schoor, C. (1995). Prestaties leerlingen in een klas lopen soms meer dan één leerjaar uiteen: interview met het PRIMA-onderzoeksteam. *Didactief*, 25 (10), 5-7.
- Slavenburg, J. H., & Peters, T. A. (1989). *Het project Onderwijs en Sociaal Milieu: een eindbalans*. Rotterdam: SAD.
- Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: a model of instructional effectiveness. *International Journal of Educational Research* 21 (2), 141-157.
- Veenman, S. A. M. (1992). Effectieve instructie volgens het directe instructiemodel. In: *Pedagogische Studiën*, 69, 242-269.
- Walberg, H. J., & Wang, M. C. (1986). Effective educational practices and provisions for individual differences. In H. J. Walberg, M. C. Wang & M. C. Reynolds (Eds.), *Handbook of special education* (pp. 113-128). Oxford: Pergamon Press.
- Willemsen, Th. F. W. P. (1994). *Remediële rekenprogramma's voor de onderbouw van de basisschool*. Academisch proefschrift. Groningen: RION/RUG.
- Wolbert, R. G. M., Span, P., & Schaap-Hummel, I. (1986). *Individualisering en differentiatie in het basisonderwijs*. Den Haag: SVO.

Auteurs

R. H. Hofman is als senior-onderzoeker verbonden aan het Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling (GION) van de Rijksuniversiteit van Groningen.

Adres: Westerhaven 15, 9718 AW Groningen.

E. B. Vonkeman is als promovendus verbonden aan de vakgroep Onderwijskunde (programma Beleid en Organisatie) van de Universiteit van Amsterdam.

Adres: Wibautstraat 4, 1091 GM Amsterdam.

Abstract

Curricula and adaptive education

R. H. Hofman & E. B. Vonkeman. Pedagogische Studiën, 1996, 73, 411-421.

Almost 50 curricula for three different subject areas were part of an extensive content analysis. The contribution of these curricula to adaptive education was examined. The results show that curricula differ considerably in the application of basic characteristics essential to adaptive education. Research on the use of curricula by teachers in age groups of 6, 7 and 8 year old pupils, shows that adaptive curricula which contain current subject contents, are more useful and less obstructing the production of adaptive education in primary schools.

Samenvatting

Ter beantwoording van de vraag of implementatie van adaptief onderwijs leidt tot betere prestaties bij technisch lezen in groep 3 van het basisonderwijs, is een quasi-experiment opgezet (untreated control group design with pretest and posttest). Er deden 456 leerlingen uit 23 basisscholen aan het onderzoek mee (12 experimentele en 11 controlescholen).

Uit het implementatie-onderzoek dat aan het quasi-experiment gekoppeld was, bleek dat de onderwijsgegenden in de experimentele conditie op een aantal kenmerken van adaptief onderwijs significant hoger scoorden dan de onderwijsgegenden uit de controlegroep. Dit betrof de volgende kenmerken: taakgerichte leertijd; het geven van directe instructie; bij het aanvankelijk lezen werken volgens de klanksynthese methodiek; en planmatig werken om uitval van leerlingen te voorkomen.

Onderwijsgegenden die dit gedrag in hoge mate vertoonden, bleken significant hogere resultaten op de Cito Drie-Minuten-Toets bij hun leerlingen te bereiken. Dit experimentele effect blijft bestaan na correctie voor zowel de voormeting als voor individuele leerlingkenmerken (intelligentie, SES en leesattitude).

1 Theoretisch kader

In de context van het beleid 'Weer Samen Naar School' (WSNS) is het begrip adaptief onderwijs opnieuw actueel (Meijer, Meijnen & Scheerens, 1993; Procesmanagement WSNS, 1994). De Stuurgroep Evaluatie WSNS (her)introduceerde het begrip in een poging het onderwerp van studie te verhelderen tegen de achtergrond van een verwarrende veelheid aan begrippen die in omloop zijn om de problematiek van opvang van leerlingen met specifieke onderwijsbehoeften aan te duiden.

In het recente verleden bestonden er uiteenlopende opvattingen over de vraag welke extra

maatregelen geschikt zijn en hoe die georganiseerd zouden moeten worden om de uitstroom van leerlingen uit het reguliere onderwijs naar het speciaal onderwijs te stoppen. Deze verschillen betreffen met name de grenzen aan de extra zorg die het reguliere onderwijs zou kunnen leveren en de verhouding tot het speciaal onderwijs. Tevens betreffen de verschillen de definiëring van de doelgroep: hebben de maatregelen alleen betrekking op de zwakke leerlingen, of op alle leerlingen? Zie voor details bijvoorbeeld: Schram (1992).

Herintroductie van het begrip adaptief onderwijs levert de volgende afbakeningen op: Door het realiseren van adaptief onderwijs houden onderwijsgegenden binnen hun reguliere onderwijs rekening met verschillen tussen leerlingen en stemmen hun onderwijs op deze verschillen af. De doelgroep betreft niet een deel van de leerlingpopulatie, maar de gehele leerlingpopulatie. Bij adaptief onderwijs wordt daarmee in eerste instantie gerefereerd aan de competenties van onderwijsgegenden en niet aan extra maatregelen en voorzieningen buiten het onderwijs in de groep, zoals speciale klassen en remedial teaching. Deze voorzieningen zijn wel belangrijk om, in het verlengde van de maatregelen die door de onderwijsgegenden binnen de groep genomen worden, op schoolniveau een continuüm van zorg te realiseren.

Deze samenhang tussen de zorg op groepsniveau en schoolniveau wordt door het Procesmanagement in het 'Basisdocument WSNS-standaards' (1994) verwoord met het thema 'Integrale leerlingenzorg'. Integrale leerlingenzorg omvat alle niveaus waarop activiteiten noodzakelijk zijn om uitstroom van leerlingen naar het speciaal onderwijs te voorkomen: de groep, de school en het samenwerkingsverband. Integrale leerlingenzorg krijgt in de eerste plaats gestalte in het onderwijsleerproces in de groep. De kern van het model voor integrale leerlingenzorg wordt met andere woorden gevormd door het primaire proces. Hiermee wordt onderstreept dat het bij WSNS niet uit-

sluitend en niet in de eerste plaats gaat om organisatorische maatregelen, maar dat het in de kern van de zaak gaat om kwaliteitsverbetering van het reguliere onderwijs. Verder krijgt integrale leerlingenzorg gestalte in maatregelen op schoolniveau. Deze maatregelen zijn op de eerste plaats bedoeld om de onderwijsgevende in staat te stellen in de groep integrale leerlingenzorg te realiseren. Ze kunnen echter tevens dienen als vangnet voor leerlingen voor wie de onderwijsgevende binnen de groep geen adequate hulp kan bieden.

Adaptief onderwijs heeft in het centrale thema integrale leerlingenzorg zijn plaats, doordat integrale leerlingenzorg op het niveau van het primaire proces gerealiseerd zou moeten worden door het geven van adaptief onderwijs.

Het grote voordeel van deze aanpak is dat aangesloten wordt bij een uitgebreid kennisbestand over de relatieve bijdrage van schoolkenmerken enerzijds en klas- en instructiekenmerken anderzijds aan het verbeteren van leerlingresultaten.

De belangrijkste factoren met verklarende waarde voor de kwaliteit van het onderwijs bestaan, blijkens diverse metastudies naar factoren van schooleffectiviteit, uit instructie- en managementfactoren op klasniveau. Ook wordt in de meta-analyses het belang van motivationele aspecten benadrukt. Deze beide elementen maken expliciet onderdeel uit van adaptief onderwijs. Naast deze kernfactoren op groepsniveau, kunnen op schoolniveau factoren onderscheiden worden die deze kernfactoren ondersteunen. Dit betreft dan vooral het creëren van voorwaarden in de organisatie voor het realiseren van goede leerlingresultaten door voortdurende sturing, ontwikkeling en professionalisering (Scheerens, 1989; Levine & Lezotte, 1990; Wang, Heartel & Walberg, 1993).

Ten behoeve van de implementatie van adaptief onderwijs zijn er in het begin van de jaren tachtig diverse programma's ontwikkeld (zie voor een overzicht Wang & Walberg, 1985). Uit vergelijking van een aantal van deze programma's is gebleken dat adaptief onderwijs geen nauwkeurig afgebakend begrip is: de programma's voor adaptief onderwijs verschillen van elkaar in de manier waarop specifieke instructiebenaderingen in het programma-

ontwerp zijn opgenomen. Ondanks de verschillen komen Wang en Lindvall (1984) en Talmage (1985) tot een opsomming van elementen die er deel van uit maken:

- een vorm van leerkrachtgestuurde instructie;
- bijhouden van de leerlingresultaten en het geven van feedback daarop;
- zelfsturing door de leerlingen wordt waar mogelijk bevorderd;
- de lerende wordt betrokken bij de planning en het management van het onderwijsleerproces.

Samengevat gaat het bij het geven van adaptief onderwijs om het verkrijgen van een gedetailleerd inzicht in het verloop van het leerproces van de leerlingen om daar gericht sturing aan te geven met als uiteindelijk doel interne sturing door de leerling zelf. Dit vereist, als vijfde punt, grondige kennis van de betreffende vakgebieden (zie Wang, 1992; Slavin e.a., 1992; Terwel, 1994). Adaptief onderwijs wordt met andere woorden gekenmerkt door: de kwaliteit van de diagnostische vaardigheid van de onderwijsgevende, het pedagogisch functioneren van de onderwijsgevende, een adequate organisatie van het onderwijsleerproces en instructievaardigheden van de onderwijsgevende, zowel in zijn algemeenheid als vakspecifiek (zie ook Houtveen & Booi, 1994). Het realiseren van adaptief onderwijs stelt derhalve eisen aan het gedrag dat de onderwijsgevende in de groep vertoont en aan de maatregelen die hij of zij treft om adaptief onderwijs in de groep te realiseren.

Er wordt niet ten onrechte op adaptief onderwijs ingezet. In diverse onderzoeken uit met name het Angelsaksische taalgebied is evidentie gevonden voor de effectiviteit van onderwijsprogramma's die voldoen aan de eerder genoemde vijf kenmerken van adaptief onderwijs bij het realiseren van goede leerlingresultaten, ook bij zwakke presteerders (Gersten & Carnine, 1984; Gersten & Keating, 1987; Walberg & Wang, 1986; Slavin, e.a., 1989; 1992; 1994; Bashi, e.a., 1990; zie voor een recent overzicht Stringfield, 1995). Behalve het programma van Wang worden deze programma's geen van alle als adaptief benoemd. Eind jaren tachtig raakte het begrip adaptief onderwijs in de Verenigde Staten enigszins in discrediet doordat het adaptieve onderwijsprogramma

van Wang inzet werd van een eerder politieke dan wetenschappelijke strijd over de mogelijkheden van volledige integratie van speciaal en regulier onderwijs (zie Hallahan, e.a., 1988; Fuchs & Fuchs, 1988a; 1988b).

In ons land is in de periode 1991-1994 een schoolverbeteringsprogramma geïmplementeerd dat als adaptief onderwijsprogramma getypeerd kan worden. Dit 'Landelijk Project Schoolverbetering' is bovendien integraal te noemen, omdat tevens implementatie van een resultaatgeoriënteerd schoolbeleid nagestreefd werd (Osinga & Boersma, 1991). Bij het realiseren van de vernieuwingen werden de scholen begeleid door onderwijsbegeleiders van zes plaatselijke dan wel regionale onderwijsbegeleidingsdiensten.

De effectiviteit van het programma is geëvalueerd (Houtveen, 1994; Houtveen, Booij & De Jong, 1995; Houtveen, Booij, De Jong & Van de Grift, 1996). Het onderzoek heeft de beperking dat uitsluitend de effectiviteit van het adaptief onderwijsgedrag van de onderwijsgeven in het onderzoek is betrokken. Het schoolniveau is buiten beschouwing gelaten. Een tweede beperking is dat het onderzoek uitsluitend betrekking had op technisch lezen in groep 3.

2 Vraagstelling en opzet van het onderzoek

Tegen de achtergrond van dit theoretisch kader luidt de vraagstelling van het onderzoek: Leidt de implementatie van adaptief onderwijs tot betere leerprestaties bij technisch lezen in groep 3 van het basisonderwijs? Om deze vraag te kunnen beantwoorden is een quasi-experiment opgezet (untreated control group design with pretest en posttest (Cook & Campbell, 1979)). Er deden 456 leerlingen uit 23 basisscholen in het onderzoek mee. Twaalf scholen werden geselecteerd om deel te nemen in de experimentele groep. Elf scholen maakten deel uit van de controlegroep. De voormeting is uitgevoerd in maart 1993 bij de groep leerlingen die het volgend schooljaar naar groep 3 zouden gaan. De nameting is uitgevoerd in juni 1994. Aan de nameting ging een implementatie-onderzoek onder de onderwijsgeven vooraf. In dit implementatie-onderzoek is vast-

gesteld of de implementatie van de experimentele variabelen in de experimentele groep in vergelijking met de controlegroep in voldoende mate gestalte kreeg.

De experimentele variabelen betreffen de realisatie van adaptief onderwijs:

- het realiseren van een hoog percentage instructietijd en een hoge mate van taakgerichtheid van de leerlingen;
- het werken volgens het model voor directe instructie;
- bij het aanvankelijk lezen gebruik maken van de principes van de klanksynthese methode en tenslotte
- het realiseren van planmatig werken om uitval van leerlingen te voorkomen.

De effectvariabele betreft bij de nameting de score op de Drie-Minuten-Toets van het Cito (Verhoeven, 1992) en bij de voormeting de score op de Auditieve-Synthesetoets uit de methode 'Alle kinderen leren lezen' (Osinga & Kooistra, 1992).

Als gevolg van de praktijksituatie waarin het onderzoek plaatsvond, kon niet voldaan worden aan de eis dat scholen random in de experimentele dan wel controle conditie geplaatst worden. De experimentele scholen werden geselecteerd uit de aan het schoolverbeteringsproject deelnemende scholen en de controlescholen zijn aselekt uit de populatie getrokken. Scholen waarvan te verwachten was dat ze gedurende de looptijd van het experiment in een fusieproces zouden geraken, zijn buiten het onderzoek gehouden. In het onderzoek zijn uiteindelijk de onderwijsgeven van groep 3 betrokken van in totaal 23 scholen, twaalf in de experimentele groep en elf in de controlegroep. Zowel de geselecteerde experimentele scholen als de controlegroep scholen zijn onderzocht op verschillen in contextkenmerken waarvan verwacht kan worden dat ze konden bijdragen aan verschillen in leerkrachtgedrag en/of leerlingresultaten. Groepsgrootte, vernieuwingservaring en het percentage allochtone leerlingen bleken in beide groepen nagenoeg gelijk.

Voor de beantwoording van de vraag of de experimentele variabelen van het project geïmplementeerd zijn, zijn er op zes momenten metingen uitgevoerd. Deze metingen hebben plaatsgevonden in de maanden januari/februari en mei/juni van 1992, 1993 en 1994. Elke me-

ting bestond uit observaties van twee lessen aanvankelijk lezen en de afname van een vragenlijst bij de onderwijsgevenden. Voor het vaststellen van de mate van werken volgens de principes van de klanksynthese methodiek vonden de metingen aan het begin van elk schooljaar plaats door middel van het invullen van logboeken door de onderwijsgevende.

3 Implementatie-onderzoek

Het implementatie-onderzoek is opgezet om een antwoord te verkrijgen op de vraag of de in het project gehanteerde begeleidingsstrategie leidt tot implementatie van de beoogde veranderingen in leerkrachtgedrag. Dit is van belang omdat het geen zin heeft om de effecten van een project te onderzoeken als niet vaststaat dat het project daadwerkelijk in de praktijk gerealiseerd is. De mate waarin de experimentele variabelen door de onderwijsgevenden zijn gerealiseerd, is gemeten met behulp van observatie-instrumenten (instructietijd, taakgerichtheid en directe instructie), een vragenlijst (planmatig werken) en logboeken (werken volgens de klanksynthese-principes). In deze paragraaf gaan we hier kort op in. Voor een uitvoerige beschrijving verwijzen we naar Houtveen, Booij en De Jong (1995).

Instructietijd en taakgerichtheid van de leerlingen

In de theoretische modellen voor het leren op school (Carroll, 1963; Bloom, 1976; Harnischfeger & Wiley, 1978) wordt de instructietijd en het efficiënt gebruik ervan door de leerlingen als een belangrijke determinant voor het leren op school beschouwd. In een groot aantal empirische onderzoeken kon het verband aangegevoerd worden tussen tijdsbesteding en leerprestaties (zie voor overzichten o.a.: Walberg, 1984; Wang, Heartel & Walberg, 1993). Derhalve is zowel de instructietijd als de taakgerichtheid van de leerlingen als variabele opgenomen.

Gedurende twintig minuten van een leesles is de hoeveelheid tijd in kaart gebracht waarin de onderwijsgevende en de leerlingen daadwerkelijk bezig zijn met activiteiten die op de leerstof gericht zijn. Dit gebeurde in observaties met behulp van een time-sampling-

instrument (Veenman, e.a., 1988, 1990). Telkens wordt de onderwijsgevende en in een tevoren vastgestelde volgorde één van de leerlingen, gedurende zeven seconden geobserveerd. Tijdens de dertien seconden daarna wordt het meest dominante gedrag gescoord. In de volgende zeven seconden wordt behalve de onderwijsgevende een andere leerling geobserveerd, enzovoorts. Op deze wijze kan met het instrument vastgesteld worden of de onderwijsgevende met inhoudelijke activiteiten (op instructie gericht) of met klasmanagement bezig is. Voor de leerlingen kan aangegeven worden of zij al dan niet met de opgedragen taak bezig zijn. Als een leerling niet-taakgericht bezig is, kan gescoord worden waaruit dit niet-taakgerichte gedrag bestaat (procedurele activiteiten, wachten, niet met de taak bezig).

Om voor elke klas een score te verkrijgen, zijn per meetmoment op verschillende tijdstippen twee lessen technisch lezen geobserveerd. In de analyses is de gemiddelde score over de twee lessen gebruikt.

Of het verschil tussen de experimentele en controlegroep scholen significant is, zijn we nagegaan door het berekenen van een *t*-toets tussen de verschillen in groeiscoringen (meting 1994 minus meting 1992) voor beide groepen.

Bij de experimentele groep nam de instructietijd toe van 63% tot 73%. Bij de controlegroep nam de instructietijd ook toe tussen de eerste en de laatste meting (respectievelijk 68% en 74%). Het verschil in groeiscoring tussen controlegroep en experimentele groep bleek voor wat betreft de instructietijd niet groot genoeg om van een betekenisvol verschil te kunnen spreken ($t = .47$).

Bij de experimentele groep was over de zes metingen een gestage groei in de taakgerichtheid van de leerlingen te constateren van 70% naar 76%, terwijl taakgerichtheid van de leerlingen in de controlegroep fluctueerde in de zes metingen. Bij de laatste meting had de controlegroep een gemiddelde taakgerichtheid van 70%. De groeiscoring in taakgerichtheid van de leerlingen bleek significant te verschillen ten gunste van de experimentele groep ($t = 3.61$).

Werken volgens het directe instructie model

Het effect van onderwijs is voor een belangrijk deel afhankelijk van de kwaliteit van de instructie. Een effectief gebleken instructiemo-

del is het model voor Directe Instructie (Rosenshine & Stevens, 1986; Rosenshine & Edmonds, 1990; Good, Grouws & Ebmeijer, 1983). Dit model is vooral effectief bij leerstof die in een stap-voor-stap benadering kan worden geleerd en die nieuw is voor leerlingen. Dit is bij technisch lezen in groep 3 het geval. De volgende te implementeren variabele wordt dan ook gevormd door de mate van werken volgens het Directe Instructie Model. Dit instructiemodel wordt zo genoemd, omdat de uitvoering van het onderwijsleerproces – zeker in de beginfase – gecontroleerd wordt door de onderwijsgevende. De kern van het model voor Directe Instructie bestaat uit de volgende hoofdfasen:

- terugblik op en activering van de voorafgaande leerstof;
- presentatie en uitleg van nieuw te leren leerstof met inbegrip van laten zien of voordoen;
- nauwgezette inoefening en begeleiding: in deze fase oefenen de leerlingen wat ze zojuist geleerd hebben en krijgen hierbij corrigerende feedback van de onderwijsgevende;
- zelfstandige of individuele oefening: in deze fase zetten de leerlingen de stap van het integreren van nieuwe kennis of vaardigheid met de reeds aanwezige kennis naar de fase van de automatisering;
- periodieke herhaling van de leerstof.

De mate van directe instructie is eveneens bepaald met behulp van observaties (zie voor het instrument: Veenman, e.a., 1988; Roelofs, 1993). Dit event-sampling instrument: de schaal 'Instructievaardigheden', bestaat uit 17 indicatoren (uitspraken), die gescoord kunnen worden op een vijf puntsschaal. In deze 17 uitspraken zijn de hoofdfasen van het model voor directe instructie geoperationaliseerd. Elk van de 17 uitspraken is gescoord op de mate van voorkomen. De score is gestandaardiseerd door de feitelijke score te delen door de maximaal mogelijke score. Hierdoor komen te scores te liggen tussen 0 en 1.

Ook bij dit instrument zijn op elk meetmoment steeds twee lessen technisch lezen geobserveerd om tot de score van dat moment te komen. De gemiddelde score hiervan vormt de score voor elke klas op het betreffende meetmoment.

Bij de onderwijsgegenden uit de experimentele groep vond in de periode tussen februari 1992 en juni 1994 een duidelijke groei van .18 op de schaal 'Instructievaardigheden' plaats. Bij de controlegroep vond in deze periode een heel beperkte groei van .05 plaats. Dit verschil in groeiscore tussen experimentele en controlegroep scholen is significant ($t = 3.68$).

Klanksynthese methode

Naast het realiseren van deze algemene instructieprincipes werd in dit schoolverbeteringsproject gestreefd naar het implementeren van instructie die erop gericht is om met name bij zwakke lezers betere prestaties te realiseren. Het gebruik van de klanksynthese methode is met name effectief gebleken voor leerlingen die moeilijk leren lezen (Dumont, 1985). In het project bleek het voor de meeste scholen niet haalbaar over te gaan op een op de klanksynthese methodiek gebaseerde methode. Als minimumeis is geformuleerd dat alle onderwijsgegenden tijdens de eerste maanden van het leesonderwijs de principes van de klanksynthese methodiek ten minste voor de zwakke lezers moeten hanteren.

Bij het werken volgens de klanksyntheseprincipes gaat het in feite om vier zaken. Het belangrijkste principe is dat in deze methode de elementaire leeshandeling centraal staat. Dat wil zeggen dat zodra er enkele letters zijn aangeleerd, met deze letters woorden worden gemaakt. Op de tweede plaats dient de onderwijsgevende volgens deze methodiek een eenduidige leesstrategie te hanteren. Eenduidig wil zeggen dat uitsluitend letters en géén woorden ingeprent worden. Bovendien mag de onderwijsgevende geen woorden aanbieden waar letters in voorkomen die niet ingeprent zijn. Een derde principe is dat vooral in het begin van het leesproces het tempo waarin letters worden aangeboden, laag dient te zijn. Het tempo dient met andere woorden afgestemd te zijn op de zwakke lezer. Ook voor de instructie in haar geheel geldt dat de zwakke lezer als norm wordt genomen in de beginperiode van het leesonderwijs. Dit vierde principe van de klanksynthese methodiek impliceert dat (vrijwel) geen enkele leerling individueel opgevangen hoeft te worden.

Om de mate vast te stellen waarin de onderwijsgegenden hun leeslessen vormgeven vol-

gens deze principes vulden de onderwijsgeven- den van groep 3 aan het begin van elk school- jaar voor tien leeslessen een registratieformul- ier in. Op elk formulier gaven ze per les aan wat het doel van de les was, welke activiteiten waren ondernomen, op welke leerlingen de ac- tiviteiten gericht waren en welke materialen hierbij gebruikt werden. Met behulp van de re- gistratieformulieren is vastgesteld of de lessen van de onderwijsgeven- den voldoen aan de bo- vengenoemde principes van het werken vol- gens de klanksynthese methodiek.

Aangezien de onderwijsgeven- den niet ver- plicht waren om volledig te gaan werken met een klanksynthese methode, dienden criteria te worden opgesteld op basis waarvan bepaald kon worden in welke mate de onderwijsgeven- den volgens de principes van deze methodiek werkten. Ten aanzien van elk van de principes zijn daarom punten toegekend afhankelijk van de mate van implementatie van het principe (zie voor een uitvoerige beschrijving: Hout- veen e.a., 1995). De punten voor elk van de principes zijn bij elkaar opgeteld en vervolgens is de score gestandaardiseerd door deze te de- len door de maximaal mogelijke score. De score kan hierdoor variëren tussen de 0 en de 1.

Over de drie schooljaren heen is er een aan- zienlijke groei van .49 opgetreden in het wer- ken met de principes van de klanksynthese methodiek in de experimentele groep. In de controlegroep vindt ook enige groei plaats (.23), deze is echter aanzienlijk geringer. Het verschil tussen de groeiscoringen in de experimen- tele en controlegroep is significant ($t = 3.00$).

Planmatig werken

In het schoolverbeteringsproject werd ernaar gestreefd uitval van leerlingen te voorkomen door systematische afstemming van het onder- wijsaanbod op de behoeften van de leerlingen: planmatig werken. De vierde te implementeren groep variabelen betreft dan ook de mate waarin onderwijsgeven- den planmatig werken bij technisch lezen. Planmatig werken omvat in essentie drie zaken. Op de eerste plaats is ken- merkend voor planmatig werken dat de extra instructie die aan de leerlingen wordt gegeven, gebaseerd is op de met behulp van toetsen vast- gelegde prestaties van de leerlingen. Op de tweede plaats is kenmerkend voor planmatig werken dat de te ondernemen (extra) activitei-

ten ten aanzien van de leerlingen worden vast- gelegd in een planningsdocument. Voor elk van de niveaus in de school (school, groep, in- dividuele leerling) zijn afzonderlijke plan- ningsdocumenten beschikbaar. Het derde ken- merk van planmatig werken is, dat er sprake is van een cyclisch verloop van de activiteiten: het afnemen van toetsen en het vastleggen van de geëigende activiteiten is geen eenmalige ge- beurtenis in een schooljaar, maar dient met een zekere regelmaat herhaald te worden. In het project is gebruik gemaakt van een model voor planmatig werken dat oorspronkelijk ontwik- keld is voor zwakke leerlingen door Kool en Van der Leij (1985). In dit model zijn proce- dures voor informatieverwerking en het opstel- len van plannen geïntegreerd. Het model be- staat uit vier fasen:

- 'Doelen stellen' bestaat uit twee subfasen: het stellen van doelen en daarbij een zekere mate van gedetailleerdheid betrachten;
- 'Signaleren van zwakke leerlingen';
- 'Problemen analyseren en leerresultaten diagnostiseren' bestaat uit twee subfasen: leerresultaten relateren aan de gegeven in- structie en resultaten van zwakke leerlingen diagnostiseren;
- 'Plannen maken en uitvoeren en resultaten evalueren' bestaat uit vier subfasen: conse- quenties trekken uit de gesignaleerde pro- blemen voor het onderwijs; activiteiten- plannen maken voor de groep als geheel; een handelingsplan maken voor de zwakke leerlingen en het bespreken van de resulta- ten van het planmatig werken in het team.

Dit model van Kool en Van der Leij (1985) vormde de basis voor de constructie van een set instrumenten voor het meten van planmatig werken. Voor elk van de negen subfasen uit het model zijn Likertschalen geconstrueerd (zie voor psychometrische gegevens: Houtveen, e.a., 1995). De scores zijn ook hier gestandaar- diseerd om een vergelijking tussen de scores op de schalen te vergemakkelijken. De scores kunnen voor elk van de schalen variëren tussen 0 en 1.

De implementatie is succesvol wanneer de experimentele groep een substantiële groei laat zien op de latere metingen op de negen schalen in vergelijking tot de eerdere metingen, terwijl de resultaten voor de controlegroep relatief sta- biel blijven op de zes metingen op elk van de negen schalen.

De experimentele groep heeft een groeiscore van in totaal 2.36 op de laatste meting op de negen variabelen, in vergelijking tot de eerste meting. De controlegroep heeft slechts een groei van in totaal .49 op de negen variabelen te zamen. Op basis hiervan kunnen we concluderen dat over het geheel genomen de scores bij de controlegroep relatief stabiel blijven. Bij acht van de negen variabelen is het verschil in groeiscore tussen de experimentele en de controlegroep significant. Op basis van deze resultaten trekken we de conclusie dat er bij de experimentele groep een substantiële groei is in planmatig werken, terwijl de controlegroep op dit punt min of meer stabiel blijft. Ook ten aanzien van deze variabele kan de implementatie van het project dus als een succes beschouwd worden.

Op basis van het bovenstaande mogen we concluderen dat we beschikken over een experimentele groep met onderwijsgevenden die op een aantal elementen van adaptief onderwijs significant hoger scoort dan de groep onderwijsgevenden uit de controlegroep. Dit maakt het zinvol om de nameting ten aanzien van de leerlingresultaten uit te voeren. Daarmee kan de vraag beantwoord worden of de onderwijsgevenden uit de experimentele groep erin slagen betere cognitieve leerlingresultaten bij technisch lezen te realiseren dan onderwijsgevenden uit de controlegroep.

4 Effectonderzoek

Bij de keuze voor een cognitieve effectvariable deed zich het probleem voor dat bij de voormeting niet met dezelfde toets gewerkt kon worden als bij de nameting, om de simpele reden dat kinderen bij het begin van groep 3 nog niet kunnen lezen. Bij de voormeting is gebruik gemaakt van de Auditieve-Synthesetoets afkomstig uit de methode 'Alle kinderen leren lezen' (Osinga & Kooistra, 1992). Met deze toets wordt de vaardigheid gemeten van de leerlingen in het samenvoegen van klanken tot betekenisvolle woorden. De toets bestaat uit 20 items. Het betreft een toets die geschikt is voor individuele afname bij leerlingen aan het eind van groep 2 of het begin van groep 3.

Voor de nameting is gebruik gemaakt van de Drie-Minuten-Toets van het Cito (Verhoe-

ven, 1992). De snelheid waarmee leerlingen woorden kunnen decoderen en ontsleutelen wordt beschouwd als een van de belangrijkste indicaties voor de leesvaardigheid. De toets bestaat uit drie leeskaarten met specifieke woordtypen. Op de eerste leeskaart staan één-lettergrepige woorden waarbij klinker en medeklinker elkaar afwisselen. Op de tweede leeskaart staan eveneens éénlettergrepige woorden, maar met een hogere moeilijkheidsgraad, doordat er meerdere medeklinkers achter elkaar in het woord voorkomen. Op de derde leeskaart tenslotte staan woorden met twee, drie en vier lettergrepen. De leerlingen krijgen voor elke kaart een minuut de tijd om zoveel mogelijk woorden goed te lezen. De score bestaat uit het aantal woorden dat de leerlingen in de gestelde tijd foutloos kunnen lezen.

Om de resultaten te kunnen corrigeren voor individuele leerlingkenmerken is tevens de intelligentie van de leerlingen, de sociaal economische achtergrond en het plezier in lezen dat kinderen ervaren, in kaart gebracht.

Voor het vaststellen van de intelligentie is gebruik gemaakt van een analogieëntest. Dit is een subtest van de non-verbale intelligentietest SON-R (Laros & Tellegen, 1991). Met deze analogieëntest wordt beoogd de vaardigheid in abstract redeneren te meten van kinderen in de leeftijd van 5½ tot 17 jaar. De subtest omvat 21 items. De homogeniteit van de toets bedroeg in het onderzoek van Laros en Tellegen .78. In ons onderzoek is opnieuw de betrouwbaarheid (Cronbachs alpha) nagegaan. Deze bedroeg .79.

De sociaal-economische achtergrond van de leerlingen is vastgesteld aan de hand van het leerlinggewicht.

Voor het meten van plezier in lezen is een leesattitudeschaal afgenomen. De gebruikte schaal is een bewerking van de Lees Attitude Schaal (LAS) voor schoolbeginners (Meijer, 1973). De schaal is ontwikkeld voor kinderen die nog niet kunnen lezen en is geschikt voor klassikale afname. De bewerkte schaal bestaat uit 19 uitspraken. Zowel bij de voor- als de nameting is de betrouwbaarheid van de nieuwe toets nagegaan door het berekenen van een Cronbachs alpha. Deze was in beide gevallen ruim voldoende (respectievelijk .83 en .86).

Het experiment is geslaagd als er bij de experimentele groep een significant hogere score

Tabel 1
Gemiddelde en standaarddeviatie van de leerlingscores

	experimentele groep		controlegroep	
	x	sd	x	sd
Drie-Minuten-Toets	93.7*	56.9	81.7*	53.3
Auditieve-Synthesetoets	15.2*	4.2	13.7*	4.6
Analogieën toets	6.3*	3.7	7.9*	4.8
Sociaal economische achtergrond	3.1	.9	3.2	1.0
Leesattitudeschaal (voormeting)	13.1	3.8	13.2	3.9
Leesattitudeschaal (nameting)	12.1	4.7	11.8	5.0

* verschil significant op het .05 niveau

perimentele groep een significant hogere score op de technisch leestoets (de Drie-Minuten-Toets) bereikt wordt dan bij de controlegroep, terwijl de overige omstandigheden, uiteraard afgezien van de experimentele conditie, gelijk blijven. De gegevens in Tabel 1 bieden een inzicht in de 'bruto' resultaten van het project.

De leerlingen uit de experimentele groep blijken 12 punten hoger te scoren op de Drie-Minuten-Toets dan de leerlingen uit de controlegroep. Dit verschil blijkt significant te zijn op het 5% niveau.

Op het eerste gezicht lijkt het experiment geslaagd. De gemiddelde score op de Drie-Minuten-Toets ligt bij de experimentele groep immers hoger dan die bij de controlegroep. In onderwijsonderzoek is het echter zelden mogelijk om een experiment uit te voeren waarbij de leerlingen voor de experimentele groep en de controlegroep volkomen random geselecteerd zijn. Niet de onderzoeker, maar de scholen bepalen immers of hun leerlingen meedoen aan een experiment. Dit brengt ook in ons experiment met zich mee dat, ondanks controle op voor de hand liggende contextvariabelen zoals groepsgrootte, vernieuwingservaring van onderwijsgegenden en het percentage allochtone leerlingen, de controlegroep en de experimentele groep niet volkomen equivalent zijn. Met niet equivalent bedoelen we, dat de controlegroep en de experimentele groep verschillende verwachte waarden hebben op ten minste één en waarschijnlijk meer variabelen (Reichardt, 1979, p. 148). Zoals uit Tabel 1 blijkt, verschillen ook in ons geval de leerlingen uit de controlegroep en de experimentele groep op diverse punten. Het is daarom heel goed mogelijk dat de verschillen op de Drie-Minuten-Toets tussen experimentele groep en controlegroep veroorzaakt worden door de verschillen in de voormeting of andere aanvangsverschillen,

zoals de intelligentie of het milieu van herkomst van de leerlingen, of het plezier dat de leerlingen in lezen ervaren.

Om na te gaan of het verschil tussen experimentele groep en controlegroep blijft bestaan, als de resultaten op de Drie-Minuten-Toets (de nameting), gecorrigeerd worden voor de resultaten op de Auditieve-Synthesetoets (de voormeting) en de controlevariabelen (intelligentie, sociaal economische achtergrond en plezier in lezen) hebben we een covariantie-analyse uitgevoerd.

Het verschil tussen experimentele groep en controlegroep op de nameting, die gecorrigeerd is voor de voormeting, de intelligentie, het sociaal milieu en het plezier in lezen van de leerlingen blijkt 16.3 te bedragen. Dit kan geïnterpreteerd worden als een verschil in het voordeel van de experimentele groep van ruim 16 punten.

De betekenis van de effecten

Voor het berekenen van de ongewogen score op de Drie-Minuten-Toets is het aantal goed gelezen woorden op elk van de drie leeskaarten van deze toets bij elkaar opgeteld. De leerlingen uit de experimentele groep hebben gemiddeld ongeveer 94 woorden goed gelezen. Voor de leerlingen uit de controlegroep waren dit ongeveer 82 woorden.

Om een goed beeld te krijgen van wat deze verschillen precies betekenen zijn we genoodzaakt terug te gaan naar de scores op de afzonderlijke leeskaarten, omdat het Cito de scores alleen voor de afzonderlijke kaarten heeft genormeerd. Het Cito heeft de toets afgenomen bij een landelijke steekproef, zodat iedere school de leerlingscores op zijn school tegen deze vergelijkingsgroep kan afzetten. Op basis van de behaalde scores zijn de leerlingen in vijf niveaugroepen verdeeld. Niveau A omvat de

Tabel 2*Vergelijkingsgegevens per kaart Cito Drie-Minuten-Toets (meetmoment groep 3/mei)*

	gemiddelde score experimentele groep	gemiddelde score controlegroep	grand mean Cito
Leeskaart 1			
A >49			
B 39-49	43.84	39.98	
C 25-38			38.40
D 18-24			
E <18			
Leeskaart 2			
A >29	30.09		
B 24-29		25.41	
C 14-23			23.42
D 9-13			
E <9			
Leeskaart 3			
A >19	19.72		
B 15-19		16.33	14.67
C 7-14			
D 4-6			
E <4			

25% hoogst scorende leerlingen, niveau B de 25% leerlingen net boven het landelijk gemiddelde en niveau C de 25% die net onder het landelijk gemiddelde scoren. De 25% laagst scorende leerlingen zijn in twee groepen ingedeeld te weten niveau D (15%) en niveau E, waarin de 10% allerzwakste leerlingen vertegenwoordigd zijn.

In ons onderzoek scoort de controlegroep gemiddeld gezien op alle drie de kaarten net boven het landelijk gemiddelde. Daarmee vormt de controlegroep in dit experiment een goed referentiepunt. Het gemiddelde van de leerlingen uit de experimentele groep ligt op alle kaarten duidelijk boven het landelijk gemiddelde. Op de eerste kaart ligt het gemiddelde halverwege niveau B. Op beide leeskaarten met woorden van een moeilijker type, ligt het gemiddelde van de leerlingen uit de experimentele groep zelfs boven percentiel 75. Overigens doet zich hierbij het verschijnsel voor dat de verdeling in de experimentele groep de verdeling in de landelijke normeringsgroep niet volgt. Relevant is tevens om te vermelden dat de leerlingen uit de experimentele groep juist op de twee moeilijkste leeskaarten significant hoger scoren dan de leerlingen uit de controlegroep.

5 Conclusie en discussie

In dit artikel zijn de resultaten vermeld van de evaluatie van het als adaptief onderwijsprogramma te typeren Landelijk Project Schoolverbetering. Uit het onderzoek komt naar voren dat er een significant verschil is op de Drie-Minuten-Toets tussen de experimentele en de controlegroep. Dit experimentele effect blijft bestaan na correctie voor zowel de voormeting als voor individuele leerlingkenmerken (intelligentie, SES en leesattitude). We kunnen dan ook concluderen dat het binnen het Landelijk Project Schoolverbetering geïmplementeerde adaptieve leerkrachtgedrag succesvol is gebleken. De resultaten lijken zonder meer een steun in de rug van het WSNS-beleid en vooral ook van de door het Procesmanagement WSNS gekozen thematiek 'Integrale Leerlingenzorg', met daarbinnen een centrale rol voor adaptief onderwijs. Kwaliteitsverbetering van het reguliere onderwijs lijkt langs deze weg mogelijk te zijn.

Daarmee is niet gezegd dat de resultaten snel en eenvoudig te realiseren waren. Integendeel, het betrof hier een veldexperiment waarbinnen aan nogal wat condities voldaan was die de resultaten mede verklaren. Tot slot van dit artikel gaan we kort op enkele van deze condities in.

Op de eerste plaats dienen kenmerken van het vernieuwingsprogramma zelf genoemd te

worden. Er was sprake van een driejarig project, waar het schoolteam in zijn geheel bij betrokken was. De school had zich voor deze periode aan het project verbonden (contract). Tevens waren er concrete afspraken gemaakt over de doelen die elke school voor zichzelf in deze periode zou willen realiseren. Een niet onbelangrijk aspect van de afspraken vormde tenslotte het feit, dat de betrokken scholen zich in de betreffende periode bepaalden tot het realiseren van deze vernieuwing. Er was met andere woorden sprake van een projectmatige aanpak, waarbij de scholen intensief werden begeleid. De betrokkenheid van externe begeleiding vormt dan ook de tweede implementatiebevorderende factor. Als derde implementatiebevorderende aanpak kan tenslotte de systematische monitoring van zowel het gedrag van de onderwijsbegeleiders als van de onderwijsgegenden genoemd worden. Systematische terugkoppeling van tussentijdse resultaten door de onderzoekers, kreeg binnen het project de functie van 'change agent'.

Literatuur

- Bashi, J., Sass, Z., Katzit, R., & Margolin, J. (1990). *Effective schools from theory and practice*. Paper presented at the ICSEI 3-5 January, Jerusalem, Israël.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Carroll, J. B. (1963). A model of school learning. In L. W. Anderson (Ed.), *Perspectives on school learning selected writing of John B. Carroll* (pp. 19-31). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation*. Design & analysis issues for field settings. Chicago: Rand McNally.
- Dumont, J. J. (1985). *Lees- en spellingsproblemen. Dyslexie, dysorthografie en woordblindheid*. Rotterdam: Lemniscaat.
- Fuchs, D., & Fuchs, L. S. (1988a). Evaluation of the adaptive learning environments model. *Exceptional Children*, 55, 115-127.
- Fuchs, D., & Fuchs, L. S. (1988b). Response to Wang and Walberg. *Exceptional Children*, 55, 138-146.
- Gersten, R., & Carnine, D. (1984). Direct instruction in mathematics: a longitudinal evaluation of low-income elementary school students. *Elementary School Journal*, 84, 395-406.
- Gersten, R., & Keating, T. (1987). Long-term benefits from direct instruction. *Educational Leadership*, 44, 28-31.
- Good, T. L., Grouws, D. A., & Ebmeier, H. (1983). *Active mathematics teaching*. New York: Longman.
- Hallahan, D. P., Keller, C. E., McKinney, J. D., Lloyd, J. W., & Bryan, T. (1988). Examining the research base of the regular education initiative: Efficacy studies and the adaptive learning environment model. *Journal of Learning Disabilities*, 21, 29-35.
- Harnischfeger, A., & Wiley, D. E. (1978). Conceptual issues in models of school learning. *Curriculum Studies*, 10, 215-231.
- Houtveen, A. A. M. (1994). Effective guidance in improving teacher behaviour: Evaluation results of the guidance strategy in the Dutch National School Improvement Project. In J. T. Voorbach (Ed.), *Teacher education*, 10. De Lier: ABC.
- Houtveen, A. A. M., & Booij, N. (1994). *Het meten van integrale leerlingzorg: Adaptief onderwijs en schoolontwikkeling*. Utrecht: ISOR, Universiteit Utrecht.
- Houtveen, A. A. M., Booij, N., & Jong, R. de (1995). *Effectieve begeleiding en leerkrachtgedrag*. Utrecht: ISOR, Universiteit Utrecht.
- Houtveen, A. A. M., Booij, N., Jong, R. de & Grift, W. van de (1996). *Effecten van adaptief onderwijs; Evaluatie van het Landelijk Project Schoolverbetering*. Alphen aan den Rijn: Samsom H. D. Tjeenk Willink.
- Kool, E., & Leij, A. van der (1985). Planmatig Handelen. In A. van der Leij (red.). *Zorgverbreding*, (pp. 69-98). Nijkerk: Intro.
- Laros, J. A., & Tellegen, P. J. (1991). *Construction and validation of the SON-R 5½-17, the Snijders-Oomen non-verbal intelligence test*. Groningen: Wolters-Noordhoff.
- Levine, D. U., & Lezotte, L. W. (1990). *Unusually effective schools: a review and analysis of research and practice*. National Center for Effective Schools Research and Development, Madison (USA).
- Meijer, C., Meijnen, G. W., & Scheerens, J. (1993). *Over wegen, schatten en sturen. Analytische beleidsevaluatie Weer Samen Naar School*. De Lier: ABC.
- Meijer, K. (1973). *De leesattitude van schoolbeginners: ontwikkeling, constructie en didaktometrische analyse van de Lees Attitude Schaal (LAS)*. Utrecht: Rijksuniversiteit Utrecht, Afdeling voor Pedagogische en Andragogische Wetenschappen, Afdeling Onderwijsresearch.

- Osinga, N., & Boersma, R.S. (1991). Schoolverbetering in Friesland. In R. Gorter & Th. Houtveen (red.), *Handboek onderwijsbegeleiding*. Alphen aan de Rijn: Samsom.
- Osinga, N., & Kooistra, A. (1992). *Alle kinderen leren lezen*. Leeuwarden: GCO.
- Procesmanagement WSNS (1994). *Basisdocument WSNS-standaards. Theoretische achtergrond en referentiekader*. Den Bosch: Procesmanagement WSNS.
- Reichardt, C. S. (1979). The statistical analyses of data from nonequivalent group designs. In T. D. Cook & D. T. Campbell, *Quasi-experimentation* (pp. 147-206). Chicago: Rand McNally.
- Roelofs, E. C. (1993). *Teamgerichte nascholing en coaching: Een experimentele studie in scholen met combinatieklassen*. Academisch proefschrift. Nijmegen: Universiteitsdrukkerij.
- Rosenshine, B. V., & Edmonds, J. (1990). New sources of improving instruction: the implicit skill study. *Journal of Personal Evaluation in Education*, 4, 59-73.
- Rosenshine, B., & R. Stevens (1986). Teacher functions. In M. C. Wittrock (Ed), *Handbook of research on teaching* (3rd ed., pp. 376-391). New York: MacMillan.
- Scheerens, J. (1989). *Wat maakt scholen effectief? Samenvatting en analyse van onderzoeksresultaten*. Den Haag: SVO.
- Schram, E. C. (1992). *Zorgverbreding in vernieuwingscholen*. Groningen: Wolters-Noordhoff.
- Slavin, R. E., Karweit, N. L., & Madden, N. A. (1989). *Effective programs for students at risk*. Needham Heights, MA: Allyn and Bacon.
- Slavin, R. E., Karweit, N. L., Wasik, B. A. (Eds.) (1994), *Preventing early school failure, research, policy and practice*. Needham Heights, MA: Allyn and Bacon.
- Slavin, R. E., Madden, N. A., Karweit, N. L., Dolan, L., Wasik, B. A. (1992). *Success for all: a relentless approach to prevention and early intervention in elementary schools*. Arlington, VA: Educational Research Service.
- Stringfield, S. (1995). Attempting to enhance students' learning through innovative programs: The case for school evolving into high reliability organizations. *School Effectiveness and School Improvement*, 6, 67-96.
- Talmage, H. (1985). What is adaptive instruction. In M. C. Wang & H. J. Walberg (Eds.), *Adapting instruction to individual differences* (pp. 113-134). Berkeley: McCutchan.
- Terwel, J. (1994). *Samen onderwijs maken. Over het ontwerpen van adaptief onderwijs*. Inaugurele rede. Groningen: Wolters-Noordhoff.
- Veenman, S., Lem, P., & Nijssen, F. (1988). *Omgaan met combinatieklassen: Een programma voor schoolverbetering*. Den Haag: SVO.
- Veenman, S., Lem, P., & Roelofs, E. (1990). Omgaan met combinatieklassen: Een onderzoek naar de effecten van een teamgericht nascholingsprogramma. *Pedagogische Studiën*, 67, 45-58.
- Verhoeven, L. (1992). *Handleiding Drie-Minuten-toets*. Arnhem: Cito.
- Walberg, H. J. (1984). Synthesis of research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp.). Washington DC: American Educational Research Association.
- Walberg, H. J. (1986). Syntheses of research on teaching. In M. C. Wittrock (Ed), *Handbook of research on teaching*, (3rd ed., pp. 214-229). New York: MacMillan.
- Walberg, H. J., & Wang, M. C. (1986). Effective educational practices and provisions for individual differences. In H. J. Walberg, M. C. Wang & M. C. Reynolds (Eds.), *Handbook of special education* (pp. 113-128). Oxford: Pergamon Press.
- Wang, M. C. (1982). *Provision of adaptive instruction: Implementation and effects*. Pittsburg: University of Pittsburg, Learning Research and Development Center.
- Wang, M. C. (1992). *Adaptive education strategies: Building on diversity*. Baltimore: Brookes Publishing Co.
- Wang, M. C., & Walberg, H. J. (Eds.) (1985). *Adapting instruction to individual differences*. Berkeley: McCutchan.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). What helps students learn? *Educational Leadership*, 52, 74-79.
- Wang, M. C., & Lindvall, C. M. (1984). Individual differences and school learning environments: Theory, research and design. In E. W. Gordon (Ed.), *Review of research in education*, Vol. XI (pp. 161-225). Washington, DC: American Educational Research Association.

Auteurs

A. A. M. Houtveen is als programmaleider onderzoek verbonden aan het ISOR van de Faculteit Sociale Wetenschappen van de Universiteit Utrecht.

N. Booij is als toegevoegd onderzoeker verbonden aan het ISOR van de Faculteit Sociale Wetenschappen van de Universiteit Utrecht.

R. de Jong is als senior onderzoeker verbonden aan het GION van de Universiteit van Groningen.

W. J. C. M. van de Grift is psycholoog en houdt zich bezig met evaluatie van het Primair Onderwijs en met name met onderzoek naar het functioneren van de schoolleiding.

Correspondentie-adres: A. A. M. Houtveen, ISOR Universiteit Utrecht, Postbus 80140, 3508 TC Utrecht.

Abstract

Adaptive instruction and pupil results

A. A. M. Houtveen, N. Booij, R. de Jong & W. J. C. M. van de Grift. *Pedagogische Studiën*, 1996, 73, 422-433.

This article reports the results of a quasi-experiment on effects of adaptive instruction on reading results of children in the first year of reading instruction in Dutch primary schools. The 456 pupils from 23 schools were involved in the research (12 experimental and 11 control group schools). Teachers in the experimental group show significantly higher adaptive instruction behavior consisting of: optimisation of time on task, use of the model for direct instruction, working according to the principles of the 'phonics construction method' for reading and degree of diagnostic teaching, than the teachers in the control group. The teachers in the experimental group proved to be successful on the effect variable: the reading results of their pupils. The difference between experimental group and control group is significant. This experimental effect does not only remain after correction for the pretest, but also after correction for intelligence, social economical background and pleasure in reading.

B. Petersen en D. Oudenhoven

Samenvatting

Als het gaat om het percentage leerlingen dat van de basisschool wordt verwezen naar het speciaal onderwijs, bestaan er in Nederland aanzienlijke verschillen tussen regio's. Uit een berekening van verwijzpercentages van OBD-gebieden in de periode 1985-1990 blijkt, dat de kans op verwijzing in die periode in het meest verwijzende gebied gemiddeld tweeënhalf maal zo groot was als in het minst verwijzende gebied. Aan de hand van een groot aantal gegevens wordt nagegaan waaraan die verschillen zijn toe te schrijven. Hieruit blijkt dat het aanbod (de capaciteit) van het speciaal onderwijs verreweg de belangrijkste determinant is. Daarentegen blijken noch demografische factoren noch kenmerken van het regionale beleid op het punt van zorgverbreding, samenwerking en verwijzing een rol van betekenis te spelen.

1 Inleiding

In een onlangs afgerond onderzoek naar de acceptatie van WSNS-beleid is de respondenten onder meer gevraagd waaraan zij regionale verschillen in de verwijzing naar het speciaal onderwijs (SO) toeschrijven: aan kenmerken van de leerling, van het milieu, van het SO-aanbod, van de basisschool of van de leerkracht? De onderzoeksgroep bestaat uit vertegenwoordigers van vijf categorieën direct en indirect betrokkenen bij de onderwijspraktijk: coördinatoren van WSNS-samenwerkingsverbanden, leden van schoolbesturen, schooldirecteuren, leerkrachten, ouders; ruim de helft van deze respondenten heeft bindingen met het SO, de overigen met het basisonderwijs. De in totaal ruim 1100 respondenten kennen gemiddeld het grootste gewicht toe aan milieukenmerken, daarna aan leerlingkenmerken en daarna aan aanbodkenmerken (Ybema, Buunk & Schripsema, 1995). In het hierna beschreven onderzoek is een soortgelijke maar meer open

vraag voorgelegd aan vertegenwoordigers van regionale onderwijsbegeleidingsdiensten (OBD's). Aan 27 respondenten die menen dat het verwijzpercentage in hun regio onder het landelijk gemiddelde ligt, is gevraagd de belangrijkste oorzaken hiervan te noemen. Verreweg het meest genoemd worden verschillende aspecten van de zorgbreedte in het basisonderwijs, op afstand gevolgd door kenmerken van het eigen OBD-beleid en van het SO-aanbod in de regio (Petersen & Oudenhoven, 1994).

Hoe realistisch zijn de veronderstellingen van de actoren uit het onderwijsveld eigenlijk? In dit artikel worden onderzoeksgegevens besproken die inzicht geven in de feitelijke effecten van deze en andere mogelijke determinanten van verwijzing, en dus ook in het realiteitsgehalte van de genoemde veronderstellingen. Dat zich in Nederland grote regionale verschillen in verwijzing voordoen, is al eerder in onderzoek aangetoond (Meijer, 1982). Vervolgens is geconstateerd dat deze verschillen bovendien tamelijk stabiel zijn: analyse van de verwijzpercentages van de 80 nodale gebieden in de jaren 1985/86, 1986/87 en 1987/88 laat zien dat er gebieden zijn die in deze drie achtereenvolgende jaren door een hoog verwijzingsniveau worden gekenmerkt en dat er gebieden zijn die door de jaren heen relatief weinig verwijzingen kennen (Oudenhoven & Bakker, 1990).

2 Probleemstelling en opzet van het onderzoek

Met het hier beschreven onderzoek wordt beoogd inzicht te bieden in de macro-factoren die de regionale verschillen in verwijzing kunnen verklaren. Met de term macro-factoren wordt hier bedoeld op factoren op gebiedsniveau, zoals de regionale onderwijsbegeleiding, het regionale voorzieningenniveau, de regionale toepassing van overheidsbeleid en de cultureel-etnische samenstelling van de regionale

bevolking. De feitelijke verwijzingsbeslissingen worden echter genomen, dan wel voorbereid, op het niveau van de (basis)scholen. Daarom is tevens informatie verzameld over de mate waarin en de wijze waarop deze macrofactoren hun invloed doen gelden op het mesoniveau, dat van de scholen.

De probleemstelling luidt als volgt:

- 1 Welke regionale factoren hebben op het niveau van OBD-gebieden invloed op (verschillen in) verwijzing van leerlingen naar het speciaal onderwijs en wat is de rol van de onderwijsbegeleidingsdiensten hierbij?
 - 2 Welke regionale factoren spelen op het niveau van de basisscholen een rol bij verwijzing?
 - 3 Welke processen op de basisscholen gaan schuil achter de geconstateerde verbanden?
- Om op deze vragen een antwoord te vinden zijn de volgende deelonderzoeken uitgevoerd:

- a een onderzoek bij onderwijsbegeleidingsdiensten: een telefonische enquête onder alle OBD's en een beschrijvende analyse van de verzamelde gegevens;
- b een causale analyse op gebiedsniveau: samenstelling van een uitgebreid databestand op gebiedsniveau, bestaande uit de onder a genoemde gegevens alsmede secundaire gegevens uit verschillende bronnen, en een causale analyse van deze gegevens;
- c een scholenonderzoek: verzameling van gegevens bij een geselecteerde groep basisscholen en een kwalitatieve analyse van deze gegevens.

Ter bepaling van het gebiedsniveau is gekozen voor een indeling van Nederland in de werkgebieden van 60 regionale onderwijsbegeleidingsdiensten (OBD's). De voor WSNS gehanteerde regio-indeling komt hier nagenoeg mee overeen.

In dit artikel zullen we ons vooral bezighouden met de eerste onderzoeksvraag, die kan worden beantwoord op grond van de resultaten van deelonderzoek b. Daarnaast zullen we kort ingaan op enkele relevante aspecten van deelonderzoek c.

Het databestand waarop de causale analyse is uitgevoerd, is samengesteld uit gegevens van uiteenlopende herkomst, die hetzij op OBD-gebiedsniveau zijn gemeten, hetzij naar dit niveau kunnen worden geaggregeerd. Behalve de

door ons verzamelde data betreffende OBD's en OBD-gebieden bevat het bestand CITO-gegevens over het prestatieniveau, CBS-gegevens over demografische variabelen, O&W-gegevens over het aanbod van SO-voorzieningen en gegevens over verwijzing, verzameld door de Nederlandse Stichting voor Statistiek (NSS).

De afhankelijke variabele is de verwijzing uit het basisonderwijs naar scholen voor (primaire) lom-, mlk-, en zmok-onderwijs. Onder verwijzing verstaan we dan gerealiseerde verwijzing, dus *plaatsing*. Deze hebben we nader geoperationaliseerd als de in een bepaalde periode in het SO-LOM, -MLK en -ZMOK geplaatste leerlingen, uitgedrukt als percentage van het gemiddeld aantal basisschoolleerlingen gedurende die periode. Deze operationalisatie kan op elke groep basisschoolleerlingen worden toegepast; in dit onderzoek gaat het in hoofdzaak om de basisschoolleerlingen van een OBD-gebied.

De door ons gehanteerde definitie sluit aan bij die van de doelgroep voor WSNS, zij het dat deze zich tot dusver heeft beperkt tot de zogenoemde groep-1 scholen (LOM en MLK). Niettemin is er veel voor te zeggen om ook de ZMOK-categorie hierin te betrekken. De grote meerderheid van de leerlingen die vanuit het basisonderwijs worden verwezen naar het speciaal onderwijs, komt terecht op scholen voor LOM, MLK en ZMOK (Ministerie O&W, 1990). Verwijzing naar andere schoolsoorten voor speciaal onderwijs, zoals scholen voor lichamelijk of zintuiglijk gehandicapten of scholen voor zeer moeilijk lerende kinderen, vindt in de meeste gevallen plaats op evident andere indicaties en vaak via een andere route. De scholen voor LOM, MLK en ZMOK zijn zowel van elkaar als van het regulier onderwijs veel minder duidelijk afgegrensd. In dit licht is het niet verrassend, dat voor de onlangs in het kader van WSNS ingestelde Regionale Verwijzingscommissies (RVC's) de ZMOK-categorie wel is 'meegenomen'.

3 Verwijzing per OBD-gebied

De van de Nederlandse Stichting voor Statistiek verkregen gegevens over het aantal verwijzingen en het totale leerlingenaantal per basis-

school per jaar hebben we met behulp van de in dat bestand opgenomen CBS-gemeentecodes geaggregeerd naar het OBD-gebiedsniveau. Met de aldus bewerkte gegevens konden we voor elk van de 60 gebieden verwijspersentages per jaar bepalen (1985/86 tot en met 1989/90). Vervolgens hebben we per gebied het gemiddelde over deze vijf schooljaren berekend. Deze 60 gemiddelden tonen aanzienlijke regionale verschillen in verwijzing. Het gemiddelde verwijsperscentage van het minst verwijzende gebied ligt voor deze vijf jaren op .38, terwijl het meest verwijzende gebied tot een gemiddelde score van .91 komt. Het landelijke verwijsperscentage over de betreffende jaren bedraagt .65, dat is één op de 150 basisschoolleerlingen per jaar. Verder blijkt nog eens dat de verwijspersentages van jaar tot jaar tamelijk stabiel zijn. De standaarddeviaties rond de gebiedsgemiddelden komen, gemiddeld over alle gebieden, uit op .07; ter vergelijking, de standaarddeviaties rond de landelijke jaargemiddelden, waarin de spreiding tussen gebieden tot uitdrukking komt, variëren van .13 tot .15. Mede gezien het vooralsnog uitblijven van WSNS-effecten is er geen aanleiding om aan te nemen dat de situatie nu, zes jaar later, zich ingrijpend heeft gewijzigd.

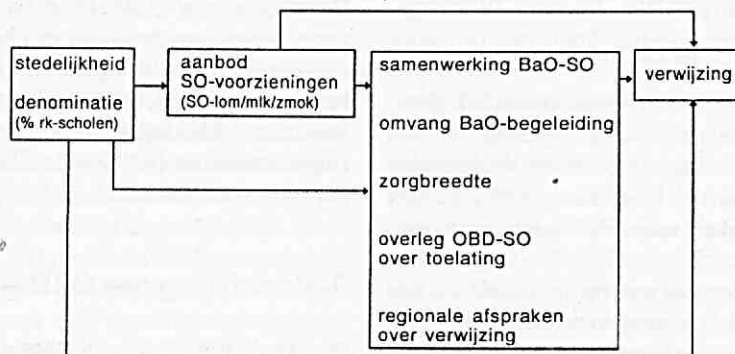
4 De causale analyse

Om te komen tot een verklaring van deze verschillen, hebben we gekozen voor het analyseren van een padmodel (d.i. een zogeheten recursief model met uitsluitend gemeten variabelen) met behulp van LISREL (Jöreskog & Sörbom, 1988). Deze procedure vereist een

keuze voor een bepaald 'beginmodel', waarin een niet al te groot aantal variabelen in een zinvolle oorzakelijke volgorde worden geplaatst. Bovendien dient een onderscheid te worden aangebracht tussen zogenoemde exogene variabelen (onafhankelijken) en endogene variabelen (afhankelijken, die ten opzichte van andere endogene variabelen ook onafhankelijken kunnen zijn).

Omdat bij de samenstelling van het databestand juist gestreefd is naar opname van zoveel mogelijk potentiële verklarende variabelen, hebben we hierin eerst stapsgewijs een sterke reductie aangebracht met behulp van correlatie-analyse, factoranalyses en multiple regressie-analyses. Op grond van theoretische overwegingen, waarin de stabiliteit van de variabelen een belangrijk criterium was, zijn de geselecteerde variabelen in een zinvolle volgorde geplaatst en is een tweetal demografische variabelen ('stedelijkheid' en 'denominatie') als exogeen aangemerkt. Het resulterende beginmodel is weergegeven in Figuur 1.

De meeste van deze variabelen spreken voor zichzelf, maar sommige behoeven een nadere toelichting. Voor de variabele stedelijkheid zijn vier indicatoren gebruikt: het aantal inwoners, de adressendichtheid, het percentage buiten Nederland geboren en het percentage eenoudergezinnen. Het aanbod van SO-voorzieningen moet worden opgevat als de (relatieve) capaciteit van de betreffende schoolsoorten en is geoperationaliseerd als het aantal leerlingen dat deze schoolsoorten bezoekt, uitgedrukt als percentage van het aantal basisschoolleerlingen op een jaarlijkse peildatum. Voor deze analyse hebben we het gemiddelde SO-aanbod over de vijf schooljaren (1985/86



Figuur 1. Het beginmodel voor de causale analyse

tot en met 1989/90) gebruikt. De samenwerking tussen basisonderwijs en speciaal onderwijs is geoperationaliseerd als de mate waarin het SO deelneemt in samenwerkingsverbanden binnen en buiten WSNS-verband (de meting is verricht in de beginperiode van WSNS). Zorgbreedte ten slotte is gemeten aan de hand van drie indicatoren: het gebruik van leerlingvolgsystemen, de aanwezigheid van Remedial Teaching en Interne Begeleiding, en de mate waarin cursussen van de OBD worden gevolgd (volgens informatie van de OBD).

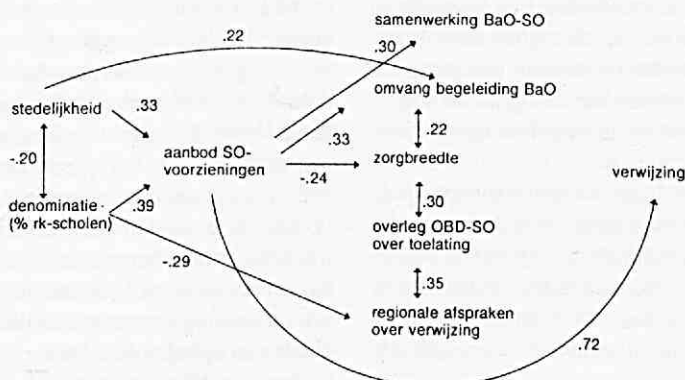
Een dergelijke analyse kan op verschillende manieren worden aangepakt. De door ons gevolgde strategie houdt in dat eerst een 'volledig model' is uitgedraaid, d.i. een model waarin alle onderlinge relaties (paden) tussen alle gekozen variabelen worden 'meegenomen'; dit betreft in bovenstaand schema alle relaties binnen elk blokje alsmede alle denkbare paden van links naar rechts. De onderlinge relaties binnen dit model zijn uitgedrukt in coëfficiënten, die de sterkte van de verschillende paden weergeven en globaal kunnen variëren van -1.00 tot 1.00, waarbij waarden in het middengebiet (rond het nulpunt) duiden op niet-bestaande of zwakke verbanden. Het 'volledige model' past, gegeven de gekozen variabelen en de gekozen causale ordening, per definitie perfect bij de data. De kwaliteit van een vereenvoudigd model is af te lezen aan een aantal maten die aangeven in hoeverre dit model afwijkt van het 'volledige model'. In Figuur 2 is een model weergegeven waarin alleen paden met een coëfficiënt boven .20 (of onder -.20) zijn opgenomen; de overige paden worden dan op nul gesteld.

We vinden hier een chi-kwadraat van 19.28 bij 22 vrijheidsgraden ($p = .63$), een prima resultaat. De AGFI (adjusted goodness of fit index) is .87, een resultaat dat als acceptabel kan worden gekenschetst. In dit model wordt 52% van de variantie in de afhankelijke variabele verklaard. Naast dit model zijn enkele andere uitgetoetst, zowel algemene als specifieke (voor lom, mlk en zmok afzonderlijk), die soms een betere fit hebben maar wat betreft de relaties tussen de variabelen niet wezenlijk afwijken van het bovenstaande model.

5 Het scholenonderzoek

Naast de causale analyse op een breed opgezet databestand is een verdiepend onderzoek verricht op twaalf basisscholen. Hiermee werd beoogd een beeld te krijgen van de mate waarin de op regionaal niveau vastgestelde verbanden ook op de scholen zichtbaar zijn en van de processen op schoolniveau die er toe bijdragen dat deze verbanden tot stand komen. Gezien de doelstelling van het scholenonderzoek en met het oog op het beperkte aantal mogelijke deelnemers, is gestreefd naar een behoorlijke geografische spreiding en een goede afspiegeling door middel van een stapsgewijze selectieprocedure (Petersen & Oudenhoven, 1994, p. 46-47). Per school hebben gemiddeld twee teamleden aan een gemeenschappelijk interview deelgenomen. De gesprekken zijn zo veel mogelijk gevoerd aan de hand van concrete verwijzingen die zich in de periode 1985-1991 hebben voorgedaan.

We kunnen de uitkomsten van deze deelstu-



Figuur 2. Padmodel voor de verklaring van regionale verschillen in verwijzing naar het LOM, MLK en ZMOK; alle paden >.20 ($n =$ maximaal 60)

die hier slechts summier weergeven. Over het algemeen hebben de geïnterviewden de indruk dat de activiteiten die in het kader van WSNS plaatsvinden zeker hun vruchten afwerpen. De opvang van zorgleerlingen in het reguliere onderwijs kan er – naar hun overtuiging – alleen maar beter van worden. Of ook het aantal verwijzingen zal afnemen, wordt in twijfel getrokken. Verder lijkt het beleid van de OBD's weinig bepalend te zijn voor de gang van zaken op de scholen rond de opvang en verwijzing van zorgleerlingen. Het beleid van de school zelf legt hier veel meer gewicht in de schaal. De al in veel onderzoek geconstateerde samenhang van de in een gebied aanwezige SO-voorzieningen met de verwijzing naar het SO wordt in grote lijnen bevestigd in de interviews. De meeste scholen denken dat de nabijheid van het SO een belangrijke factor is, maar achten deze niet zo zeer bepalend voor het schoolbeleid als wel voor de houding van de ouders. Over het algemeen is de weerstand van ouders tegen plaatsing van hun kind in het SO – eenmaal overtuigd van de noodzaak ervan – minder groot wanneer de SO-school op niet al te grote afstand van huis of de huidige school ligt. Vrijwel alle geïnterviewden hebben de indruk dat het ontbreken van een SO-school in hetzelfde dorp (bijna) altijd betekent dat – onder invloed van de ouders – minder snel wordt verwezen. Eenzelfde verband bestaat ook andersom (veel SO-voorzieningen betekent veel verwijzingen); of het even sterk geldt wordt door sommigen betwijfeld. Bij de houding van de ouders ten aanzien van verwijzing, spelen vermoedelijk meer nog dan de geografische nabijheid – die in de praktijk vaak redelijk goed te overbruggen is – de vertrouwdsheid en bekendheid met de expertise van de school (die door de nabijheid zeker worden bevorderd) een grote rol. Voor sommige ouders kan de signatuur van de SO-school eveneens drempelverlagend werken.

Geconcludeerd kan worden dat zich op de meeste scholen processen voordoen – of die zich nu wel of niet hoofdzakelijk bij de ouders afspelen – die inderdaad kunnen leiden tot een sterke samenhang tussen het aantal verwijzingen en het voorzieningenniveau in een gebied.

6 Conclusies en discussie

Uit onze analyses op regionaal niveau over de periode 1985- 1990 blijkt dat het aanbod van speciaal onderwijs een zeer krachtig, rechtstreeks effect heeft op de verwijzing naar het speciaal onderwijs. Uit het scholenonderzoek komt naar voren dat de ouders mede verantwoordelijk zijn voor dit mechanisme: zij moeten uiteindelijk ja of nee zeggen, en zij zeggen nee wanneer zij de afstand tussen huis en school voor hun kind te groot vinden. Andere wezenlijke effecten op verwijzing zijn op regionale schaal niet vastgesteld.

In gebieden waar relatief veel aan *zorgverbreding* wordt gedaan (in de zin van specifieke faciliteiten en deskundigheidsbevordering), blijkt niet minder te zijn verwezen dan in gebieden waar relatief weinig aan zorgverbreding wordt gedaan. Hieruit zijn niet zonder meer conclusies te trekken ten aanzien van de haalbaarheid van het WSNS-beleid, om een drietal redenen. Ten eerste gaat het door ons onderzochte tijdvak (1985-1990) nog vooraf aan de start van WSNS (1992), waarin een intensivering van zorgverbredingsprocessen was voorzien. Ten tweede komt de door ons gebruikte operationalisering van het begrip zorgverbreding niet geheel overeen met die welke de laatste tijd in WSNS-kader het meest wordt gehanteerd, waarbij de nadruk wordt gelegd op het zogeheten adaptief onderwijs. Ten derde weten we niet zeker in hoeverre de geconstateerde processen zich in de veronderstelde volgorde hebben voorgedaan, omdat een deel van de variabelen gelijktijdig is gemeten en een ander deel via een retrospectieve vraagstelling.

Al met al roepen onze bevindingen omtrent zorgverbreding op regionale schaal diverse vragen op die ten dele door het 'beleid' en ten dele door onderzoek zouden kunnen worden beantwoord. Is zorgverbreding bedoeld als een instrument om de verwijzing omlaag te brengen of is het uitsluitend een middel om de gevolgen op te vangen van stabilisatie en mogelijk reductie van het speciaal onderwijs? Dient het accent in het zorgverbredingsbeleid te liggen op het omgaan met verschillen in de klas, in plaats van op schoolfaciliteiten? Heeft het ontbreken van effecten op verwijzing wellicht te maken met het feit dat de tot nu toe gerealiseerde zorgbreedte overal nog ontoereikend is?

Of met het feit dat deze effecten zich pas na jaren openbaren?

In het recente WSNS-beleid is het al langer bestaande zorgverbredingsbeleid opgenomen als één onderdeel van de strategie. Daarnaast heeft WSNS nog andere ijzers in het vuur, zoals intensivering van de *samenwerking* tussen regulier en speciaal onderwijs. Weliswaar is de samenwerking in het kader van WSNS pas de laatste jaren op grote schaal van de grond gekomen, maar het geeft te denken dat in gebieden waar het speciaal onderwijs al langer deelnam aan samenwerkingsverbanden, niet minder wordt verwezen dan in gebieden waar deze samenwerkingsverbanden pas in WSNS-kader zijn gerealiseerd. Naar aanleiding hiervan kan men soortgelijke vragen stellen als ten aanzien van de zorgverbreding. Overigens zijn er verschillende onderzoeksresultaten die wijzen op een mogelijk 'averechts effect', zowel van samenwerking als van zorgverbredingsinspanningen: beide processen kunnen in de praktijk leiden tot een verhoogde gevoeligheid voor specifieke leerlingproblematiek en daardoor de neiging tot verwijzen gedurende korte of langere tijd doen toenemen; Meijer (1988) spreekt in dit verband van de 'sensitiviteitsthese' (zie verder: Van IJendoorn, 1991; Meijer, Meijnen & Scheerens, 1993; Van Rijswijk, Kool & Van Beukering, 1993).

Basisscholen verwijzen niet veel en niet zo maar. Niet veel, omdat per school per jaar slechts één of enkele leerlingen naar het speciaal onderwijs worden verwezen. Niet zo maar, omdat scholen pas verwijzing adviseren na een tamelijk langdurige en behoedzame procedure, waarin ook alternatieve oplossingen worden overwogen en uitgetoet. Het mag dan ook geen verbazing wekken, dat verwijzing op schoolniveau niet als een probleem wordt ervaren, eerder als de oplossing voor een probleem. Als op basisscholen een appèl wordt gedaan om de groei van het speciaal onderwijs tot staan te brengen, zullen de meeste individuele scholen zich dan ook niet aangesproken voelen, aangezien die groei op schoolniveau meestal niet zichtbaar is. Waar sprake is van intensieve regionale samenwerking en informatie-uitwisseling, lijkt een dergelijk appèl een betere kans van slagen te hebben.

In de inleiding tot dit artikel zijn enkele in het onderwijsveld levende *denkbeelden* over

regionale verschillen in verwijzing weergegeven. We kunnen op grond van de hiervoor besproken resultaten concluderen dat deze denkbeelden niet erg realistisch zijn. Met name de veronderstelling dat regionale verschillen in verwijzing te verklaren zijn uit verschillen in problematiek, vindt geen steun in de door ons geanalyseerde gegevens. Er is geen substantieel verband gevonden tussen het verwijzingspercentage en verschillende indicatoren van achterstand (inclusief CITO-toetsgemiddelden). Wel is het mogelijk dat er in het verleden aanvankelijk een dergelijke oorzaak-gevolgrelatie is geweest, waarna de tot stand gekomen voorzieningenstructuur ervoor gezorgd heeft dat het betreffende niveau van verwijzing tot op heden is geconsolideerd. Maar ook ten aanzien van de totstandkoming van de voorzieningenstructuur zijn er indicaties voor een alternatieve hypothese. De aanwezigheid van een relatief hoog voorzieningenniveau, gekoppeld aan een relatief hoog verwijzingspercentage, in de overwegend rooms-katholieke gebieden wijst bij voorbeeld meer in de richting van een cultuurhistorische verklaring (Petersen & Oudenhoven, 1994, p. 65-66).

Dit alles betekent dat onze bevindingen wel onderbouwing geven aan een beleid dat is gericht op *verevening* (toekenning van gelden voor speciale zorg aan regio's naar rato van het aantal basisschoolleerlingen, los van de omvang van de bestaande SO-voorzieningen). Zoals uit voorgaande conclusies blijkt, komen enkele andere veronderstellingen van het WSNS-beleid er vooralsnog minder gunstig van af.

Tot voor kort is wellicht te gemakkelijk aangenomen dat een gericht beleid van regionale samenwerking en zorgverbreding wel het beoogde effect (verminderde instroom in het SO) zou opleveren. Naarmate een structureel effect langer uitblijft, lijkt het onderwijsbeleid meer geneigd tot het nemen van globale aanvullende maatregelen, in de hoop dat deze het gewenste effect alsnog bewerkstelligen. Maar ook deze maatregelen (instelling Regionale Verwijzingscommissies, nieuw systeem van leerlingenzorg, invoering van klasse-assistenten) missen een solide empirische onderbouwing. Omdat de in dit en ander grootschalig onderzoek beschikbare gegevens slechts een beperkt

inzicht kunnen bieden in de werking van bepaalde maatregelen, verdient het aanbeveling specifieke vormen van onderzoek op te zetten om het effect van zulke maatregelen vast te stellen. Een geschikt middel lijkt het opzetten van beleidsexperimenten, waarbij een bepaalde maatregel in één of twee gebieden wordt ingevoerd, terwijl andere beleidsaspecten zoveel mogelijk stabiel worden gehouden. Door de ontwikkeling van het verwijsperscentage in zo'n gebied gedurende enkele jaren te volgen en te vergelijken met de trend in vergelijkbare gebieden en met de landelijke trend, verkrijgt men meer zekerheid omtrent het effect van zulke maatregelen.

Literatuur

- IJzendoorn, W. J. E. van (1991). *(Neven-)Effecten van formatiefixatie*. Groningen: RION.
- Jöreskog, K. G. & Sörbom, D. (1988). *LISREL VII: A guide to the program and applications*. Chicago: SPSS Inc.
- Meijer, C. J. W. (1982). *LOM-onderwijs: Verwijzing en toelating*. Groningen: RION.
- Meijer, C. J. W. (1988). *Verwijzing gewogen: een studie naar de determinanten van verwijzing naar speciaal onderwijs*. Groningen: RION.
- Meijer, C., Meijnen, W. & Scheerens, J. (1993). *Over wegen, schatten en sturen. Analytische beleids-evaluatie Weer Samen Naar School*. De Lier: Academisch Boeken Centrum.
- Ministerie van Onderwijs en Wetenschappen (1990). *Weer samen naar school. Perspectief om leerlingen ook in reguliere scholen onderwijs op maat te bieden*. Zoetermeer.
- Oudenhoven, D. A. J. M., & Bakker, J. T. A. (1990, t.b.v. SVO-programmering 1991): *Oriënterende analyse van regionale cijfers met betrekking tot de verwijzing van leerlingen naar het speciaal onderwijs*. Interne notitie.
- Petersen, B., & Oudenhoven, D. (1994). *Verwijzing naar het speciaal onderwijs: een analyse van regionale verschillen*. Nijmegen: ITS.
- Rijswijk, C. M. van, Kool, E., & Beukering, J. T. E. van (1993). *Weer samen naar de Amsterdamse basisschool*. Amsterdam: GPI.
- Ybema, J. F., Buunk, A. P., & Schripsema, R. D. (1995). *Streefbeeld en verevening. Acceptatie van wijzigingen in de financiering van het speciaal onderwijs*. De Lier: Academisch Boeken Centrum.

Auteurs

B. Petersen en D. Oudenhoven zijn als onderzoekers verbonden aan het Instituut voor Toegepaste Sociale Wetenschappen van de Katholieke Universiteit Nijmegen.

Adres: ITS, Postbus 9048, 6500 KJ Nijmegen.

Abstract

How do regional differences in special education placement come about?

B. Petersen en D. Oudenhoven. *Pedagogische Studiën*, 1996, 73, 434-440.

With regard to the proportion of pupils placed in special schools in the Netherlands there are considerable differences between regions. To find out how these regional differences come about a causal analysis was performed on a large database. The supply (i.e. capacity) of special education proved to be the decisive factor. Neither demographic variables nor characteristics of regional education policy turned out to be relevant.

C. J. W. Meijer

Samenvatting

In deze bijdrage geven we een beeld van de gedachten die de kern vormen van het WSNS-beleid. We schetsen daarbij de doel-middel ketenstructuur die ten grondslag ligt aan het WSNS-beleid. Vervolgens beschrijven we de stand van zaken van het WSNS-beleid. Interessant is verder het internationale perspectief. De lessen die daaruit zijn te trekken in het kader van WSNS worden dan ook kort besproken. Van groot belang is verder de vraag of er in het veld al vooruitgang wordt geboekt. We sluiten af met een vooruitblik. Wat leren de gegevens ons over de mate waarin WSNS in de toekomst succes zal kunnen boeken?

1 WSNS: doelen en middelen

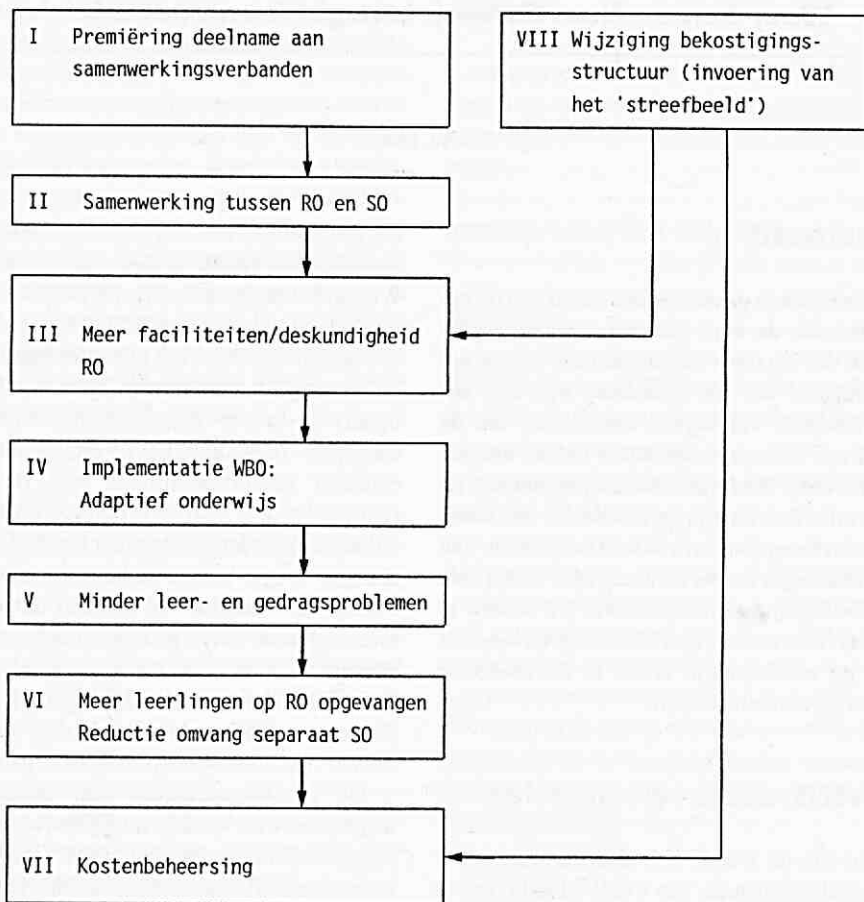
In één van de WSNS beleidsdocumenten, *3 x Akkoord* (Ministerie van O&W, 1991), wordt als doelstelling van het WSNS-beleid geformuleerd dat het erom gaat om binnen het reguliere onderwijs de opvang van meer leerlingen mogelijk te maken (*3 x Akkoord*, p. 3). Door het verhogen van de deskundigheid van leraren en de samenwerking met enkele soorten speciaal onderwijs zou de zorg in de basisschool moeten worden verbreed. De toenmalige staatssecretaris Wallage en de (besturen)organisaties benadrukken in *3 x Akkoord* dat de huidige systeemscheiding vervangen dient te worden door een continuüm van zorg binnen samenwerkingsverbanden van scholen voor regulier en speciaal onderwijs. Hierbij wordt het begrip 'continuüm van zorg' gedefinieerd in termen van een samenhang tussen regulier en speciaal onderwijs in bestuurlijk, organisatorisch en onderwijskundig opzicht, zodanig dat leerlingen met specifieke pedagogisch-didactische behoeften op passende en flexibele wijze begeleid kunnen worden, terwijl fixering van plaats en aanbod wordt teruggedrongen (p. 11). Men ziet een dergelijke doelstelling overigens niet

als iets nieuws; het zou gaan om een verdere specificering van de doelen zoals vervat in de Wet op het Basisonderwijs (WBO).

De kerngedachte van WSNS is dan als volgt te beschrijven. Door het bijeenbrengen van regulier en speciaal onderwijs (systeemscheiding opheffen) wordt het basisonderwijs beter toegerust (deskundigheid, faciliteiten), om conform de doelstellingen van de WBO (zorgverbreding, realisatie ononderbroken ontwikkeling) eerder en een meer flexibel georganiseerde zorg te bieden leidend tot een vermindering van schooluitval, hetgeen moet leiden tot een afname van de groei van het speciaal onderwijs zodat er in de toekomst sprake is van financiële beheersbaarheid (Meijer, Meijnen & Scheerens, 1993, p.41). De doelen van Weer Samen Naar School zijn daarmee geschetst.

De middelenstructuur kan uiteengelegd worden in twee 'hoofd'-middelen, te weten de samenwerkingsverbanden (SWV's) en de voorgenomen overheveling van de meerkosten van het onderwijs voor kinderen met leer- en opvoedingsmoeilijkheden (LOM), voor moeilijk lerende kinderen (MLK) en voor in hun ontwikkeling bedreigde kleuters (IOBK), naar de aan de SWV's deelnemende basisscholen (in de Kamer aangeduid als het 'streefbeeld'). In Figuur 1 worden de doelen en middelen schematisch weergegeven.

De twee middelen, SWV's en 'streefbeeld' worden beide als voorwaarden gezien voor het tot stand brengen van een continuüm van zorg. Volgens Meijer, Meijnen en Scheerens is het eerste middel (SWV) weer een voorwaarde voor het tweede middel (andere bekostigingsstructuur). De SWV-oprichting kan als zodanig worden geïnterpreteerd als de organisatorische voorbereiding voor de invoering van een nieuwe bekostiging waarbij middelen van het speciaal onderwijs (SO) worden overgeheveld naar het reguliere onderwijs (RO).



Figuur 1. Centrale causale keten WSNS-beleid (bron Meijer, Meijnen & Scheerens, 1993)

2 Recent WSNS-beleid

Recent hebben de Akkoordpartners in beginsel overeenstemming bereikt over een nieuw stelsel voor de financiering van speciale zorg in het onderwijs. In de Tweede Kamerbehandeling van oktober 1995 is dat akkoord bekrachtigd. Een belangrijk deel van de middelen die nu direct toevloeien naar de scholen voor speciaal onderwijs (althans voor wat betreft de zogenaamde groep-1 scholen) zullen, na een geleidelijk overgangstraject, worden overgeheveld naar de scholen voor basisonderwijs. De scholen voor speciaal onderwijs kunnen dan nog een genormeerd bedrag tegemoet zien op basis van een deelnamecijfer van 2% (het landelijk gemiddelde deelnamecijfer aan de groep-1 scholen is nu 3,8%). De overige middelen (op basis van de resterende 1,8%) gaan dan naar de basisscholen. Deze scholen overleggen vervol-

gens in het samenwerkingsverband (SWV) hoe de middelen zullen worden besteed.

Daarmee is sprake van een gedeeltelijke invoering van het oorspronkelijke streefbeeld waarin werd geanticipeerd op een volledige overheveling van zorgmiddelen. Verder is er sprake van een vereveningsprincipe. In de toekomst krijgen alle SWV's een zorgbudget op basis van het totaal aantal leerlingen. De verevening heeft een 'dubbel' karakter: zowel de financiering van speciale zorg van de in de SWV's samenwerkende basisscholen, als de directe bekostiging van de speciale voorziening wordt verevend en niet meer gebaseerd op het aantal gediagnosticeerde en/of toegelaten leerlingen, maar op het totaal aantal leerlingen in een gebied. Deze regeling zal vanaf 1998 geleidelijk worden ingevoerd met een overgangsperiode van vier jaar om te komen tot het gewenste einddoel.

Vanaf 1998 zijn de middelen 'samenwerkingsverbanden' en 'verandering bekostiging' min of meer operationeel en dan kan dan ook worden nagegaan of de doelen inderdaad worden bereikt. Niettemin worden aan de oprichting van SWV's ook immanente zorgverbredende effecten toebedacht. Zo wordt van de oprichting van SWV's verwacht dat scholen kennis zullen nemen van elkaars werkwijzen en deskundigheden, hetgeen een adequate zorg voor 'probleemleerlingen' in het reguliere stelsel zal faciliteren. De oprichting van de SWV's en de daaraan verbonden middelen vormen in deze gedachtengang de korte termijn middelenstructuur en de invoering van de nieuwe bekostiging de lange termijn middelenstructuur.

Cruciaal is natuurlijk de vraag of op grond van de beschikbare kennis verwacht kan worden dat met de inzet van deze middelen de WSNS-doelen inderdaad worden bereikt. Meijer, Meijnen en Scheerens (1993) waren daar pessimistisch over. Zij stelden dat het op grond van de onderzoeksresultaten niet aannemelijk is dat de oprichting van samenwerkingsverbanden sec leidt tot de realisatie van een van de doelstellingen van de WBO: adaptief onderwijs. Een belangrijk argument is volgens hen dat de samenwerking niet wordt gevoed vanuit een structureel ingebouwde of intrinsieke noodzaak. Waar het om gaat is het inbouwen van de noodzaak tot structurele samenwerking. Een van die manieren kan de invoering van de nieuwe bekostigingswijze zijn. Immers, de verdeling van geoormerkte middelen kan een noodzaak tot structurele samenwerking impliceren, waar alle betrokkenen de relevantie van inzien. Een verandering van de bekostiging is ook om andere redenen zeer geboden. De huidige financiering kent een aantal duidelijke nadelen. Immers, waar aan de ene kant zorgverbreding en differentiatie worden nagestreefd, is er aan de andere kant ook sprake van mechanismen in de regelgeving die de groei van het SO bevorderen. In de praktijk betekent dit dat het toelaten van leerlingen in het SO wordt beloond, het terugplaatsen niet (voldoende) wordt bevorderd en ook het handhaven van leerlingen met problemen in het reguliere stelsel niet wordt gestimuleerd. Het vigerende 'tweesparen'-systeem beloont dus segregatie en ontmoedigt integratie. De ervaringen met het WSNS-proces wijzen erop dat veel scholen

strategisch gedrag vertonen (Meijer, Peschar & Scheerens, 1995). De wijze van bekostiging draagt mede bij aan het feit dat speciale scholen steeds weer nieuwe wegen zoeken om de financiën op peil te houden. Het lijkt er sterk op dat een overheidsbeleid dat de financieringsstructuur niet aanpast weinig succesvol kan zijn en zelfs bijdraagt aan de groei.

Het blijft vanzelfsprekend nog de vraag of de doelstellingen van de WBO bij een gewijzigde bekostiging wel gerealiseerd worden. Daarover kan men kort zijn. Het realiseren van adaptief onderwijs voor een (grotere) verscheidenheid aan leerlingen kan nooit een direct gevolg zijn van de oprichting van samenwerkingsverbanden. Ook niet als deze inhoudelijk echt samenwerken. Daar is veel meer voor nodig: men denke aan geschikte methoden (zie de bijdrage van Hofman & Vonkeman in dit nummer), aan leerkrachten die bereid zijn verschillen tussen leerlingen te accepteren en dat als uitgangspunt voor hun onderwijs te beschouwen en aan andere voorwaarden op school en bovenschools niveau (zie de bijdrage van Houtveen, Booij, De Jong en Van de Grift in dit nummer).

Er is dus veel meer nodig dan SWV-vorming en een verandering van de bekostiging. Maar vastgesteld moet worden dat zonder een vorm van bovenschoolse samenwerking en de noodzakelijke faciliteiten (die omgezet kunnen worden in 'resources' als extra deskundigheid/tijd/aandacht) een meer adaptief en retentief onderwijsconcept (zie Doornbos, 1991) weinig kans van slagen heeft. In dit opzicht vormt de invoering van een nieuwe bekostiging één van de noodzakelijke voorwaarden voor het creëren van een gewijzigd onderwijssysteem dat (ook) ten goede komt aan leerlingen met speciale pedagogisch-didactische behoeften. En voor de invoering van een andere bekostiging is de oprichting van SWV's weer een noodzakelijke voorwaarde. Immers, het niet organisatorisch op elkaar betrekken van speciaal en regulier onderwijs zou niet alleen tot een te grote en oncontroleerbare verdamping van de beschikbare middelen en deskundigheden leiden, het zou ook de zorg voor leerlingen met specifieke onderwijskundige behoeften bedreigen.

Vanuit deze gezichtspunten is SWV-vorming als een noodzakelijke (maar niet voldoende) voorwaarde te definiëren voor het op

termijn verkrijgen van adaptief en retentief onderwijs. De invoering van een nieuwe bekostiging is een logische en noodzakelijke volgende stap. Zonder een dergelijk bindend element kan men van samenwerking niet veel effecten verwachten.

Concluderend kunnen we stellen dat op voorhand de effectiviteit van het WSNS-beleid niet te hoog mag worden ingeschat. In feite is de meest reële verwachting dat er in het veld niet veel inhoudelijke veranderingen te constateren zullen zijn.

3 Enkele buitenlandse lessen

Uit buitenlandse gegevens (zie UNESCO, 1994; OECD, 1995) blijkt dat het bij het streven naar integratie niet gaat om een aantal simpele aanpassingen van het vigerende onderwijssysteem, maar om verdergaande ingrepen. Het standaard curriculum, de uniforme klas, de sterk summatieve beoordelingsprocedures en zittenblijven worden als conflicterend beschouwd met de wens om verschillen tussen leerlingen als uitgangspunt voor het onderwijs te hanteren. Wet- en regelgeving voor regulier en speciaal onderwijs zullen moeten worden geïntegreerd. Ook is een sterk gescheiden opleidingstraject niet in het belang van het streven naar integratie. Alle leerkrachten zouden vooral moeten worden opgeleid in het hanteren van verschillen tussen leerlingen. Het speciaal onderwijs krijgt in veel landen een nieuwe rol: het gaat functioneren als een (collegiale) hulpbron voor de reguliere scholen. Speciale scholen richten zich daarbij minder op het geven van onderwijs aan leerlingen met problemen, maar zijn vooral georiënteerd op de ondersteuning van leerkrachten in het basisscholen. Het vergaren en verspreiden van kennis, de ontwikkeling van materialen, de voorlichting van ouders, de opleiding van leerkrachten en dergelijke zijn de nieuwe taken van de speciale centra.

Decentralisatie wordt als een belangrijke voorwaarde^o voor integratie gezien (Pijl & Meijer, 1994). Verantwoordelijkheid op het gedecentraliseerde niveau kan bijdragen aan het succes van het integratiestreven. Er is verder brede ondersteuning voor de gedachte dat wettelijke en met name financiële regelingen in

de pas moeten lopen met het integratie-streven. De 'incentive-structuur' zal zodanig gericht moeten zijn dat integratie er door bevorderd wordt.

In buitenlandse studies wordt de bekostiging mede verantwoordelijk gehouden voor het instandhouden van een gescheiden stelsel van regulier en speciaal onderwijs (OECD, 1995). Er is in een aantal landen en internationale organisaties dan ook veel aandacht voor de bekostigingssystematiek en de effecten daarvan op de zorg voor leerlingen met specifieke behoeften (UNESCO, 1994; OECD, 1995). In het algemeen ziet men veel nadelen in een aanbodgerichte financiering. Als alternatief wordt genoemd: een budgetfinanciering van de regio of van samenwerkende scholen. Met dit systeem is in een aantal landen enige ervaring opgedaan (Canada bijvoorbeeld) en in het algemeen lijkt er ondersteuning te zijn voor het (geoordeelde) budgetfinancieringsmodel (Porter, 1994). Dit model zou meer recht doen aan de behoefte om lokale/regionale oplossingen te creëren die toegesneden zijn op de betreffende situatie.

Nederland is op weg om een aantal veranderingen in het onderwijs door te voeren die kunnen bijdragen aan het realiseren van meer geïntegreerd onderwijs. Maar er moet nog veel gebeuren. Het speciale stelsel is de keerzijde van het sterke uniforme basisonderwijs waar het omgaan met verschillen tussen leerlingen nog niet van de grond is gekomen. Het systeem selecteert sterk en laat dat op verschillende karakteristieke punten zien. Verder is de regelgeving teveel gericht op segregatie en dit leidt mede tot de leidende positie die Nederland in vergelijking met andere landen inneemt in het plaatsen van leerlingen op scholen voor speciaal onderwijs.

4 Bedreigingen en kansen

Een recente studie naar de stand van zaken van WSNS laat zien dat WSNS een proces is van vallen en opstaan. Kwantitatieve gegevens wijzen nog niet op een stabilisatie van het absolute aantal leerlingen in de groep-1 scholen, laat staan op een inkrimping van de omvang van het speciale stelsel.

In het veld zijn er verder interessante ont-

wikkelingen waar te nemen. Op grote schaal is een taakdifferentiatie in de basisscholen doorgevoerd die er in veel scholen toe heeft geleid dat er naast een directeur ook een interne begeleider in de basisschool werkzaam is. Hoewel informatie ontbreekt over de precieze werkwijze en de mogelijke effecten ervan, kan deze ontwikkeling toch positief worden geduid. Een persoon die verantwoordelijk is voor de speciale zorg, lijkt immers een belangrijke voorwaarde voor het in gang brengen van veranderingen. Een tweede positieve waarneming is dat er sprake is van een aantal projecten die pionieren met een bepaald integratiemodel. De betekenis van dergelijke voortrekkers kan nauwelijks worden overschat. Succeservaringen kunnen een enorme stuwende functie hebben.

De beschrijving van de ontwikkelingen in de praktijk laat ook zien dat er sprake is van een aantal gevaren. Op dit moment zijn de samenwerkingsverbanden voornamelijk bezig met de organisatorische invulling van het WSNS-proces, in plaats van met de inhoudelijke vernieuwing: het streven naar adaptief onderwijs. Een organisatorisch accent bij aanvang van een dergelijke vernieuwing is echter begrijpelijk. Wel zal voortdurend gewezen moeten worden op de noodzaak van inhoudelijke veranderingen op het niveau van het primaire proces. Dat de invoering van adaptief onderwijs tot aardige effecten kan leiden, is door Houtveen, Booi, De Jong en Van de Grift (dit nummer) voldeende gedemonstreerd. Goede methoden, waarbij rekening wordt gehouden met verschillen tussen leerlingen, zijn daarbij een belangrijke voorwaarde. Methoden verschillen zeer in de mate waarin deze adaptief onderwijs mogelijk maken (zie Hofman & Vonkeman in dit nummer). Dat leerkrachten adaptieve methoden als meer bruikbaar beoordelen voor het bereiken van adaptief onderwijs, is dan ook bemoedigend.

In het voorgaande is betoogd dat van het WSNS-beleid zoals dat tot nu toe vorm heeft gekregen in feite nog niet veel resultaten zijn te verwachten. De stand van zaken in het veld lijkt die 'voorspelling' te bekrachtigen. De samenwerking is lang nog niet overal inhoudelijk van aard en het primaire proces staat zeker nog niet centraal. Door het naar voren halen van de oorspronkelijk voor 1997 geplande discussie over

de eventuele verandering van de bekostiging (in de Tweede Kamer bloemrijk omschreven als het 'of en hoe' van het 'streefbeeld') is echter belangrijke tijdswinst geboekt. Verschillende grote veranderingen staan nu op stapel, zoals de invoering van één wet voor het primair onderwijs in 1998, de gedeeltelijke overheveling van de middelen voor speciale zorg van de speciale scholen naar de samenwerkende basisscholen en de verevening van de zorgmiddelen.

De vraag is natuurlijk of hiermee de doelen van het WSNS-beleid meer kans van slagen hebben. Op grond van internationale gegevens en de analytische doordenking van de centrale doel/middel-relaties kan worden gesteld dat de kans daarop inderdaad hoger moeten worden ingeschat dan vanuit de huidige situatie mogelijk is. En hoewel daarvoor nog veel meer nodig is, is een aantal belangrijke stappen gezet die voorwaardenscheppend zijn voor de verdere ontwikkeling van een geïntegreerde speciale zorg. De empirische evaluatie zal moeten uitwijzen of de WSNS-doelstelling realistisch is.

Literatuur

- Doornbos, K. (red.). (1991). *Samen naar school. Aangepast onderwijs in gewone scholen*. Nijkerk: Intro.
- Meijer, C.J.W., Meijnen, G.W., & Scheerens, J. (1993). *Over wegen, schatten en sturen. Analytische beleidsevaluatie 'Weer Samen Naar School'*. De Lier: Academisch Boeken Centrum.
- Meijer, C.J.W., Peschar, J.L., & Scheerens, J. (1995). *Prikkels*. De Lier: Academisch Boeken Centrum, 1995.
- Ministerie van Onderwijs & Wetenschappen (1991). *3x Akkoord*. 's-Gravenhage: SDU.
- OECD (1995). *Integrating students with special needs into mainstream schools*. Paris: OECD.
- Pijl, S.J., & Meijer, C.J.W. (1994). *Integratie in internationaal perspectief. Een studie naar bronnen voor het regulier onderwijs*. Groningen: RION.
- Porter, G. (1994). *Organization of schooling: achieving access and quality through inclusion*. Paper presented at the world conference on special educational needs: access and quality, 7-10 June, Salamanca, Spain.
- UNESCO (1994). *World conference on special needs education: access and quality. Final report*. Paris: UNESCO.

Auteur

C.J.W. Meijer Coördinator Evaluatie Weer Samen Naar School.

Adres: Zilverschoon 25, 9801 LM Zuidhorn.

Abstract

Together to school again: review and perspective

C.J.W. Meijer. Pedagogische Studiën, 1996, 73, 441-446.

In The Netherlands the perspective of integrating special and regular schools is defined by the concept of 'adaptive education'. This concept forms the educational core of recent government policy on integration. The goal is to offer special provision for more children within regular schools through various measures. By transferring the facilities and resources of special education more readily to regular schools and by having schools working together in schoolclusters, the government is seriously attempting to achieve integration. A first step towards this was to establish regional school clusters of both regular and special schools. There are also plans to change financial regulations whereby the ultimate aim is to transfer parts of special education funding to regular schools.

M. Verkuyten, B. Kinket en Ch. van der Wielen

Samenvatting

In dit artikel wordt ingegaan op de vraag wat kinderen tussen de tien en dertien jaar verstaan onder discriminatie en of er gedeelde voorstellingen zijn over diverse aspecten van discriminatie. Het onderzoek is uitgevoerd onder 193 autochtone en allochtone leerlingen van vijf basisscholen. Voor deze leerlingen is het prototypisch voorbeeld van discriminatie de situatie waarin een autochtoon kind een allochtoon kind uitscheldt zonder aanvaardbare reden. Situaties van oneerlijk verdelen en sociale uitsluiting worden in mindere mate als discriminatie beoordeeld. Er zijn weinig verschillen tussen autochtone en allochtone leerlingen en er zijn geen verschillen voor sekse en leeftijd. De resultaten worden geïnterpreteerd in termen van het idee van sociale representaties.

Inleiding

Bestaande, vooral Engelstalige, studies naar discriminatie onder kinderen richten zich op vragen zoals de mate waarin discriminatie voorkomt, de achtergronden van discriminatie en de reacties van de slachtoffers van discriminatie (zie Aboud, 1988; Clark, 1966; Milner, 1983). Vooroordeel en discriminatie bij kinderen worden bestudeerd door de voorkeur en afkeer van kinderen vast te stellen of door reacties ten opzichte van kinderen van een andere etnische groep te observeren. Wat nauwelijks tot niet bestudeerd wordt is de betekenis van begrippen zoals vooroordeel en discriminatie voor kinderen zelf. Vooroordeel en discriminatie zijn wetenschappelijk-technische concepten die door de onderzoeker gedefinieerd en geoperationaliseerd worden. Maar deze begrippen hebben ook een alledaagse betekenis en figureren in het denken over situaties en inter-etnische relaties. Saharso (1992) en Leeman (1994) hebben dit laten zien voor adolescenten die in Nederland wonen. Ook voor

oudere kinderen heeft met name het begrip 'discriminatie' waarschijnlijk een bekende klank. Dit begrip wordt veelvuldig gebruikt in de media en maakt veelal ook onderdeel uit van het curriculum in scholen die zich bijvoorbeeld met multi- en intercultureel onderwijs bezig houden. In dit artikel gaan wij in op de vraag wat oudere kinderen onder discriminatie verstaan.

Onderzoek naar deze vraag is om ten minste twee redenen van belang. Ten eerste is het mogelijk dat bestaande onderzoeken niet op adequate wijze de betekenissen en kennis van kinderen bestuderen. Het is genoegzaam bekend dat het type vragen dat aan kinderen gesteld wordt van invloed is op hun antwoorden. Onderzoekers hebben controle over het type vragen waardoor zij niet alleen specifieke antwoorden kunnen uitlokken, maar ook voorbij kunnen gaan aan de betekenissen van kinderen zelf. Dit laatste kan ook het geval zijn in de onderwijspraktijk wanneer opvattingen over discriminatie van leerkrachten onvoldoende aansluiten bij die van kinderen. Daarom is het nodig om aandacht te schenken aan de wijze waarop kinderen bepaalde gedragingen zien en beoordelen.

Ten tweede heeft een dergelijk onderzoek theoretische relevantie. Het meeste onderzoek onder kinderen naar onderwerpen die te maken hebben met 'ras' en 'etniciteit' gaan uit van ofwel sociale leertheorieën ofwel cognitieve ontwikkelingsmodellen. Sociale leertheorieën leggen de nadruk op processen van identificatie en het mechanisme van belonen-straffen alsmede op het belang van de ouders, vrienden en de media bij het ontwikkelen van vooroordelen en discriminerend gedrag. Cognitieve ontwikkelingsmodellen baseren zich op het werk van Piaget en Kohlberg. Zo gebruikt Aboud (1988) de theorie van Piaget om de ontwikkeling van vooroordeel bij kinderen te verklaren. Maar deze ontwikkelingsmodellen kennen een aantal beperkende uitgangspunten. In de volgende paragraaf zal worden betoogd dat voor onder-

zoek naar de betekenissen van discriminatie het idee van sociale representaties een vruchtbare aanvulling is op bestaande benaderingen.

1 Sociale representaties

Onder de invloed van Piaget is vanaf het midden van de jaren zestig een constructivistisch perspectief gaan domineren in de ontwikkelingspsychologie. Verschillende onderzoekers hebben het werk van Piaget naar cognitieve ontwikkeling uitgebreid naar het terrein van sociale cognitie waarbij werd verondersteld dat kinderen interacteren met anderen in antwoord op 'dis-equilibrium' (Damon, 1977; Turiel, 1983). Nu zijn er vanzelfsprekend verschillen tussen onderzoekers die werken vanuit een constructivistisch perspectief, maar er zijn ook overeenkomsten (Emler, 1987; Emler & Ohana, 1993). De constructivistische benadering legt de nadruk op de actieve rol van het kind. Kinderen worden verondersteld kennis te verwerven door zelf informatie uit de omgeving te interpreteren en te gebruiken. De nadruk ligt op de vermogens van het individuele kind en zijn actieve verhouding tot de omgeving. De aandacht gaat uit naar abstracte processen van denken en het oordelen wordt gezien als een proces waarbij feitelijke informatie en ervaringen worden geïnterpreteerd aan de hand van abstracte principes. Zo zouden morele oordelen - zoals het oordeel van discriminatie - gebaseerd zijn op twee componenten: 1) de kenmerken van een bepaald geval zoals die worden begrepen en 2) het algemene morele principe dat wordt toegepast op deze kenmerken. Het theoretische resultaat is een nadruk op onderliggende principes met een sequentie van stappen of fasen in de ontwikkeling van aspecten van cognitie. Met iedere fase worden de cognitieve processen logisch consistent en meer aangepast aan de werkelijkheid. Deze ontwikkeling zou zelfregulerend zijn omdat het een consequentie is van de spontane cognitieve activiteit van het kind die gestuurd wordt door het streven naar (hoger) cognitief evenwicht. Het is het kind dat probeert betekenis te geven aan de wereld en die zijn of haar kennis construeert.

De constructivistische benadering is door verschillende auteurs bekritiseerd. Alterna-

tieve begrippen zoals 'situated cognition', 'social language' en 'socially shared knowledge' zijn voorgesteld (zie Resnick, Levine & Teasley, 1991). Gemeenschappelijk hieraan is het idee dat de sociale omgeving waarin cognitieve processen plaatsvinden een integraal onderdeel is van deze processen en niet alleen de omringende context ervan. Ook is er het verwijt dat vorm en inhoud van het denken te sterk worden gescheiden en dat de aandacht eenzijdig uitgaat naar het eerste ten koste van het tweede. De alternatieve benaderingen benadrukken dat theorieën van sociale cognitie niet gevangen moeten blijven in de individualistische en proces doctrine van de constructivistische benadering (Corsaro, 1992). Cognitie wordt gezien als ingebed in een historische, culturele en sociale context. Mensen die deel uitmaken van dezelfde cultuur of groep zouden in hun denken gedeelde kennis en opvattingen gebruiken die mede van invloed zijn op de wijze van denken. Cognities zijn in deze visie geen individuele constructies die ieder kind voor zichzelf moet maken, maar sterk beïnvloed door de opvattingen en ideeën die gangbaar zijn in de sociale omgeving van het kind: betekenisgeving is een sociaal proces en betekenissen zijn sociale producten.

Een theoretisch raamwerk dat eveneens de nadruk legt op het sociaal gedeelde karakter van cognitie is het idee van sociale representaties dat door Moscovici (1982, 1984) is geïntroduceerd binnen de Europese sociale psychologie. In de voetsporen van Durkheim bestudeert Moscovici het collectieve, gedeelde karakter van voorstellingen waarvoor hij het begrip 'sociale representaties' gebruikt. Sociale representaties zijn onderling samenhangende voorstellingen, ideeën en waarden die mensen in gemeenschap met elkaar maken en veranderen (zie Zeegers & Jansz, 1988). Ze zijn het resultaat van een sociaal proces van betekenisgeving en overstijgen daarmee de individuele persoon. Sociale representaties gaan over alledaagse kennis, over hoe mensen kennis delen en met elkaar hun sociale realiteit constitueren. Representaties vormen de bouwstenen van onze gedeelde werkelijkheid en dringen zich aan ons op als zijnde de werkelijkheid. Ze zijn niet sociaal omdat het representaties van de sociale wereld zijn en ook niet zozeer omdat ze sociaal van oorsprong zijn, maar vooral omdat

ze door mensen gedeeld worden en de werkelijkheid constitueren.

Verscheidende empirische studies hebben het idee van sociale representatie als uitgangspunt genomen voor onderzoek naar het denken van kinderen (Duveen & Lloyd, 1990; Emler & Dickinson, 1985; Emler, Ohana & Moscovici, 1987; Lloyd, 1987; Uzzell & Blud, 1993). In dit artikel gaan wij om een drietal redenen uit van dit idee. Ten eerste gaat het bij sociale representaties om alledaagse betekenissen en de inhoud van het denken. Dit sluit aan bij onze doelstelling om na te gaan wat oudere kinderen inhoudelijk onder discriminatie verstaan. Nagegaan zal worden of kinderen concepties van discriminatie delen. We zullen de veronderstelling onderzoeken dat er sociaal gedeelde voorstellingen zijn over diverse aspecten van discriminatie, zoals om wat voor soort gedrag het gaat bij discriminatie, wie er bij discriminatie betrokken is en waarom discriminatie voorkomt. Daarnaast besteden we aandacht aan de morele redeneringen van kinderen door na te gaan waarom kinderen een specifieke handeling als al dan niet discriminatie beoordelen. Deze verklaringen geven inzicht in de sociaal geaccepteerde redenen om gedragingen te classificeren en daarmee te rechtvaardigen.

Ten tweede gaat constructivistisch onderzoek vooral uit van een probleem-oplossend model van cognitie, waarbij de nadruk ligt op cognitieve structuren en processen. In tegenstelling hiermee wordt met het idee van sociale representatie de inhoudelijke kant benadrukt. Beoordelingen zouden gebaseerd zijn op betekenissen die gedeeld worden. In aanvulling op de kenmerken van een concreet geval en de algemene morele principes die hierop worden toegepast hebben morele oordelen met andere woorden ook te maken met relevante betekenissen die leden van een groepering delen. Morele oordelen zijn ingebed in sociale representaties. Eén van de implicaties is dat kinderen afhankelijk van de gangbare opvattingen over discriminatie specifieke gevallen als discriminatoir zullen beoordelen. Een van deze opvattingen is dat discriminatie gerelateerd is aan maatschappelijke posities. Discriminatie wordt typisch gezien als een handeling van leden van de dominante meerderheidsgroep of autochtonen ten opzichte van minderheden of allochtonen. Dit gangbare idee wordt uitgedra-

gen door de media en ook in scholen. Zo heeft Leeman (1994) laten zien dat in lessen over inter-etnisch samenleven impliciet de boodschap wordt overgedragen dat autochtonen domineren en allochtonen slachtoffer zijn. Hierbij is de achterliggende doelstelling om autochtonen discriminatie af te leren en om allochtonen te leren zich weerbaar op te stellen. In ons onderzoek hebben we aan kinderen korte verhaaltjes gepresenteerd waarin de rollen van dader en slachtoffer van discriminatie systematisch gevarieerd werden voor etniciteit. Er waren verhaaltjes met een autochtoon kind in de rol van dader en een allochtoon respectievelijk autochtoon kind in de rol van slachtoffer. En er waren verhaaltjes met een allochtoon kind als de dader en een autochtoon of allochtoon kind als slachtoffer. Als de morele beoordeling van discriminatie voornamelijk afhangt van de toepassing van formele morele principes – zoals discriminatie is selectief en ongerechtvaardigd gedrag op basis van een groepskenmerk – kan verwacht worden dat de vier verhaaltjes overeenkomstig zullen worden beoordeeld in termen van discriminatie. Als daarentegen kennis in de vorm van opvattingen en ideeën eveneens een rol spelen dan kunnen we verwachten dat de situatie waarin een autochtoon kind de dader is en het allochtone kind het slachtoffer eerder als discriminatie zal worden gezien.

Ten derde stelt de cognitieve ontwikkelingsbenadering dat de verwerving van sociale kennis in essentie een individueel cognitieve aangelegenheid is waarbij afhankelijk van de fase van ontwikkeling informatie en ervaringen worden verwerkt. Het behoren tot een sociale meerderheids- of minderheidsgroep kan in dit verband van belang zijn. In het algemeen bekleden leden van etnische minderheden een lagere positie in de samenleving dan autochtonen en lopen ze een grotere kans om geconfronteerd te worden met vooroordelen, stigmatisering en uitsluiting. Dit is niet alleen feitelijk zo, maar wordt ook zo gezien en ervaren door allochtonen jongeren (Verkuyten, 1992). Saharso (1992) laat zien dat dit besef en eigen ervaringen een effect hebben op wat onder discriminatie wordt begrepen. Autochtone en allochtone kinderen kunnen hierdoor gedeeltelijk anders over discriminatie denken. Een dergelijk verschil wordt ook verondersteld vanuit het idee van sociale representaties, maar alleen

als autochtone en allochtone kinderen in een verschillende sociale omgeving leven. Als beide groepen in eenzelfde omgeving leven dan zal er juist overeenstemming zijn in sociale representaties en daarmee nauwelijks verschil tussen autochtone en allochtone kinderen. In dat geval zullen voorstellingen, opvattingen en ideeën gedeeld en gebruikt worden om eigen gedragingen en die van anderen van betekenis te voorzien. In overeenstemming met deze gedachte constateerden Verkuyten en Masson (1993) geen verschillen in etnisch categoriseren en in het gebruik van etnische indelingen tussen autochtone en allochtone kinderen van dezelfde basisscholen. School is een belangrijk onderdeel van de omgeving van kinderen en is vaak gerelateerd aan andere belangrijke kenmerken zoals buurt, opleiding en beroep van de ouders en sociale status. De autochtone en allochtone kinderen die aan ons onderzoek hebben meegewerkt zaten op dezelfde scholen die voornamelijk door leerlingen uit de lokale buurt worden bezocht. In navolging van het idee van sociale representaties is onze voorstelling dan ook dat er tussen autochtone en allochtone kinderen weinig verschillen zullen zijn in de betekenis van discriminatie.

Een laatste opmerking is op zijn plaats en betreft de status van het idee van sociale representaties. Moscovici spreekt uitdrukkelijk van de sociale representatie theorie. Maar over het gebruik van het begrip 'theorie' is heel wat onenigheid. Dit is mede het gevolg van verschillen in opvatting over wat een wetenschappelijke theorie is. Diverse auteurs hebben kritiek geuit op Moscovici en spreken liever van het idee van sociale representaties in plaats van van een theorie. Het is hier niet de plaats om nader op deze discussie in te gaan (zie De Rosa, 1994). Van belang is om er op te wijzen dat Moscovici eerder een heuristisch raamwerk biedt dan een theorie in de zin van een samenhangend stelsel van proposities waaruit voorstellingen worden afgeleid die kunnen worden getoetst. Onderzoek naar sociale representaties is voornamelijk beschrijvend van aard en tracht de verschillende beelden en opvattingen over een bepaald onderwerp in kaart te brengen. In dit artikel pretenderen we dan ook niet – zo dit al mogelijk is – om (onderdelen) van de sociale representatie 'theorie' te toetsen. Wij gebruiken het idee van sociale representaties eerder

als uitgangspunt en richtingwijzer bij het formuleren van verwachtingen over wat kinderen onder discriminatie zullen verstaan en bij het interpreteren van de resultaten.

2 Methode

2.1 Onderzoeksgroep

Het onderzoek is uitgevoerd op vijf multi-etnische basisscholen in Eindhoven. In totaal hebben 193 leerlingen aan het onderzoek meegewerkt uit de groepen 7 en 8. Het percentage autochtone leerlingen in de onderzoeksgroep is 53. Van de 47% allochtone leerlingen hebben er 50 een Turkse vader en moeder, 14 een Marokkaanse vader en moeder en zijn er 16 van Surinaamse afkomst. De meerderheid van de allochtone leerlingen had een Islamitische achtergrond. Het percentage jongens in de onderzoeksgroep is 52 en er was geen significant verschil tussen de etnische groepen voor sekse (χ^2 -kwadraat (1,190) = 3.5, $p > .05$). De gemiddelde leeftijd van de leerlingen was 11,4 jaar en er was een verschil in leeftijd voor etniciteit (χ^2 -kwadraat (3,191) = 14.1, $p < .001$). De allochtone leerlingen waren gemiddeld iets jonger (11.2) dan de allochtone leerlingen (11.6). De gegevens zijn klassikaal verzameld met behulp van een vragenlijst met open en gesloten vragen.

2.2 Materiaalverzameling

Vooraf zijn twee voorstudies uitgevoerd om na te gaan welke soorten gedrag kinderen als vormen van discriminatie beschouwen. In een eerste voorstudie onder dertien leerlingen is gevraagd of zij wisten wat discriminatie is ('ja-nee'). Vervolgens werd aan de leerlingen gevraagd om schriftelijk uit te leggen wat discriminatie is. Hieruit kwam naar voren dat kinderen vooral drie soorten gedrag als discriminatie zagen: 1) schelden en pesten, 2) oneerlijk verdelen en 3) uitsluiten bij spelen. In een tweede voorstudie zijn deze drie soorten gedrag vertaald in twaalf korte verhaaltjes over concrete voorvallen tussen twee kinderen. Deze verhaaltjes zijn voorgelegd aan 21 leerlingen. Op grond van hun antwoorden en opmerkingen zijn de verhaaltjes aangepast voor de hoofdstudie.

In het eigenlijke onderzoek is een vragen-

lijst bestaande uit twee delen gebruikt. In het eerste deel is eenzelfde vraag als in de eerste voorstudie voorgelegd. De leerlingen werd gevraagd of ze wisten wat discriminatie is. Vervolgens werd aan hen gevraagd om een kort verhaaltje te schrijven over discriminatie, bij voorkeur iets dat ze zelf hadden meegemaakt of gezien. Deze verhaaltjes zijn geanalyseerd op een drietal kenmerken: 1) de aard van het gedrag, 2) wie er betrokken is bij het discriminerende gedrag en 3) waarom het discriminerende gedrag zich voordoet.

In het tweede deel kregen de leerlingen twaalf korte verhaaltjes voorgelegd over concrete voorvallen tussen twee kinderen. Er waren steeds vier verhaaltjes voor drie typen gedrag: schelden, oneerlijk verdelen en uitsluiten. De vier verhaaltjes over schelden hadden te maken met etnisch schelden en met het onderscheid tussen stad en platteland (bijv. 'Naïd loopt in de pauze buiten. Hij komt uit een klein dorp en is pas verhuisd naar de stad. Naïd zit bij Idir in de klas. Als Idir vlak langs Naïd loopt roept Idir stomme koeienmester tegen hem'). De vier verhaaltjes over oneerlijk verdelen gingen over het verdelen van snoep en speelgoed (bijv. 'Petra is jarig en mag in de klas trakteren. Ze heeft een grote zak minirepen en mag iedereen er twee geven. Wanneer ze bij Güllü komt geeft ze haar er maar één in plaats van twee'). De verhaaltjes over uitsluiten beschreven gebeurtenissen waarbij een kind wilde meedoen, maar door een ander kind werd geweigerd (bijv. 'Buiten spelen een paar kinderen tikkertje op het schoolplein. Ivo vraagt aan Robert of hij mee mag doen. Robert wil dit niet en laat Ivo niet meedoen'). Voor ieder type gedrag was de rol van dader en slachtoffer systematisch gevarieerd naar etniciteit (zie Tabel 1). Het onderscheid tussen autochtone en allochtone dader en slachtoffer werd in de verhaaltjes geoperationaliseerd door het gebruik van bekende Nederlandse namen (zoals Maarten, Petra en Susan) en bekende Islamitische namen (zoals Fatima, Omar en Salim). In alle verhaaltjes waren de twee kinderen van dezelfde sekse en zes verhaaltjes gingen over twee meisjes en de andere zes over twee jongens.

De leerlingen werd bij ieder verhaaltje als eerste gevraagd of er volgens hen sprake was van discriminatie. Vervolgens werd gevraagd om het antwoord schriftelijk kort toe te lichten.

Deze toelichtingen zijn geanalyseerd naar redenen die de leerlingen geven om het gedrag al dan niet als discriminatie te beoordelen. Deze analyse is apart gedaan voor de bevestigende en de ontkennende antwoorden. Om de verschillende redenen te ordenen zijn de verklaringen gelezen en herlezen om tot een complete inventarisatie te komen van de gebruikte redenen. De volgende stap betrof een verdere reductie van de diverse redenen en het geven van een naam aan de geconstrueerde categorieën. Hiertoe zijn frequenties gebruikt waarbij een reden in het coderingsstelsel werd opgenomen als het door ten minste 10% van de kinderen op één van de twaalf verhaaltjes werd gebruikt. Voor de bevestigende antwoorden heeft dit geresulteerd in zeven categorieën. De eerste drie verwijzen naar abstracte morele waarden die criteria vormen om gedrag als discriminatie te beoordelen: 1) *Moraliteit*, verwijzend naar morele evaluaties in het algemeen (zoals 'Dat is niet goed', 'Dat mag je niet zomaar doen' en 'Dat hoort toch niet'). 2) *Verdelende rechtvaardigheid* waarbij werd benadrukt dat je eerlijk moet delen en dat iedereen gelijk moet worden behandeld, 3) *Gemeenschappelijkheid*, waarbij er op werd gewezen dat niemand zo maar of bij voorbaat mag worden uitgesloten. De vierde categorie had betrekking op de beschikbaarheid van (4) *Gedragsalternatieven*. Als er volgens de leerlingen alternatieven beschikbaar waren dan werd het gedrag als discriminerend beoordeeld ('Hij hoeft niet zo te doen, hij kan toch ook weggaan'). De twee volgende categorieën verwijzen naar kenmerken van het slachtoffer. 5) *Onverantwoordelijk*, waarbij er op werd gewezen dat het slachtoffer er niets aan kan doen dat hij of zij anders is of doet zoals hij of zij doet. 6) *Etniciteit*, waarbij er op werd gewezen dat het om discriminatie gaat omdat het slachtoffer onheus wordt bejegend of behandeld op grond van zijn of haar etnische achtergrond. De zevende categorie (7) is een *rest-categorie* waarin bijvoorbeeld antwoorden zoals emotionele identificatie ('Ik zou het ook niet leuk vinden') gecodeerd zijn. Tenslotte zijn in een aparte categorie de antwoorden gecodeerd waarin geen verklaring werd gegeven, maar bijvoorbeeld alleen de gebeurtenis in het verhaaltje werd herhaald.

Voor de ontkennende antwoorden zijn vijf categorieën gebruikt bij het coderen. 1) *Vrij-*

heid, verwijzend naar antwoorden waarbij werd gesteld dat iedereen het recht heeft om te doen wat men wil en te spelen met wie men wil. 2) *Verantwoordelijkheid*, waarbij het voorval niet als discriminatie werd beoordeeld omdat het slachtoffer gedeeltelijk verantwoordelijk werd geacht voor het gedrag van de dader ('Misschien heeft hij eerst gescholden'). 3) *Verzachting*, waarbij werd benadrukt dat het voorval niet ernstig is of dat de dader het niet zo bedoelde of meende ('Hij plaagde waarschijnlijk alleen maar'). 4) *Etniciteit*, waarbij werd aangegeven dat het niet om discriminatie gaat omdat het twee kinderen van dezelfde etnische groep betreft of een allochtone dader en autochtone slachtoffer. 5) *Rest-categorie*, zoals antwoorden dat iedereen wel scheldt en dat het daarom geen discriminatie is. Tenslotte zijn ook hier de antwoorden waarin geen verklaring werd gegeven apart gecodeerd.

Omdat de meeste kinderen (95%) naar één reden verwezen is de verklaring van ieder kind slechts één keer gecodeerd. Het coderen is gedaan door een van de auteurs. Vervolgens is een dertigtal toevalsgewijs geselecteerde vragenlijsten gecodeerd door een andere auteur. De mate van codeerovereenkomst uitgedrukt in Cohens kappa was 0,84. Aangezien het in ons onderzoek om nominale data gaat zijn bij de statistische analyses voornamelijk niet-parametrische toetsen uitgevoerd (chi-kwadraat). Hierbij is ook gekeken naar de sterkte van de samenhang (contingentie-coëfficiënten, *cc*). Gezien de diversiteit aan etnische groepen is alleen een vergelijking gemaakt tussen autochtone en allochtone leerlingen gezamenlijk.

3 Resultaten

3.1 Wat is discriminatie?

Tweënnegentig procent van de leerlingen zegt te weten wat discriminatie is ($N = 177$). Hierbij is er een significant verschil voor etniciteit (chi-kwadraat $(1,193) = 12.3, p < .001$). Van de autochtone leerlingen geeft 98% aan te weten wat discriminatie is terwijl dit percentage voor de allochtone leerlingen 84% is. Er waren geen significante verschillen voor sexe (chi-kwadraat $(1,189) = 3.0, p > .05$) en leeftijd (chi-kwadraat $(3,190) = 1.8, p > .10$). In de navolgende analyses zijn alleen die leerlingen be-

trokken die aangeven te weten wat discriminatie is. We zullen verder ook niet ingaan op mogelijke verschillen naar sekse en leeftijd omdat er bij de verschillende analyses geen verschillen tussen jongens en meisjes en tussen jongere (10-11 jaar) en oudere kinderen (12-13 jaar) naar voren zijn gekomen.

De eerste vraag is *wat* voor gedrag de kinderen beschrijven op de vraag om een concreet voorbeeld van discriminatie te geven. Voor 67% van de kinderen is schelden een vorm van discriminatie. Tien procent beschrijft voorbeelden van oneerlijk verdelen en 8% van sociale uitsluiting. Verder beschrijft 5% gevallen van fysiek geweld en de overige 10% beschrijven andere vormen van gedrag, zoals pesten. Vergelijkingen tussen autochtone en allochtone leerlingen geeft geen significant verschil te zien (chi-kwadraat $(4,177) = 4.8, p > .10, cc = 0.21$).

De tweede vraag is *wie* er bij het discriminerende gedrag betrokken is. Discriminatie wordt door de kinderen voornamelijk gezien als een handeling van een persoon ten aanzien van een andere persoon. Negenenvertig procent van de kinderen beschrijft discriminatie in deze termen en 31% beschrijft een situatie waarbij een persoon wordt geconfronteerd met een groep. Er is geen significant verschil tussen autochtone en allochtone leerlingen (chi-kwadraat $(1,177) = 3.7, p > .10, cc = 0.08$). Een tweede kenmerk dat in de beschrijving naar voren komt is de leeftijd van slachtoffer en dader. Negen procent beschrijft een volwassene in de rol van dader en 51% een kind als dader. De resterende 40% gebruikt leeftijd onspecifieke aanduidingen zoals 'je' en 'iemand'. Er is in dit verband een verschil voor etniciteit (chi-kwadraat $(2,140) = 9.1, p < .01, cc = 0.25$). Allochtone leerlingen beschrijven vaker dan autochtone leerlingen een volwassene in de rol van dader (23% tegen 4%) en minder vaak een kind (26% tegen 57%). Kinderen worden beschreven als de slachtoffers van discriminatie door 65% van de leerlingen, 32% gebruikte leeftijd onspecifieke termen en 3% beschrijft een volwassene als slachtoffer. Er zijn hierbij geen verschillen tussen autochtone en allochtone leerlingen (chi-kwadraat $(2,177) = 1.9, p > .10, cc = 0.11$). Een derde kenmerk is de etnische achtergrond van de dader en slachtoffer van discriminatie. In het geval van de dader

geeft 75% van de kinderen geen informatie over de etnische achtergrond. Van de leerlingen die hierover wel informatie geven beschrijft 72% een autochtoon als dader ('Nederlander', 'witte') en 28% een allochtoon ('buitenlander', 'bruine'). Voor de totale percentages is er geen verschil tussen autochtone en allochtone leerlingen (chi-kwadraat (2,177) = 3.9, $p > .10$, $cc = 0.16$). Tweeëntwintig procent van de leerlingen geeft geen informatie over de etnische achtergrond van het slachtoffer. Van de leerlingen die hierover wel informatie geven beschrijft 16% een autochtoon slachtoffer en 84% een allochtoon slachtoffer. Opvallend hierbij is dat Turkse kinderen relatief vaak (27%) als slachtoffer worden beschreven. Er is een verschil voor etniciteit omdat 25% van de autochtone leerlingen een autochtone slachtoffer beschrijven terwijl dit percentage 4% is onder de allochtone leerlingen (chi-kwadraat (2,177) = 17.6 $p < .01$, $cc = 0.26$). Een vierde kenmerk, tenslotte, is de vraag hoe de leerlingen hun eigen rol beschrijven in het concrete voorbeeld van discriminatie dat ze hadden meegemaakt of gezien. Zes procent beschrijft zichzelf als dader, 71% als slachtoffer en 23% als toeschouwer. Er is een tendens voor allochtone leerlingen om in vergelijking met autochtone leerlingen zichzelf meer in de rol van slachtoffer te beschrijven (81% tegenover 62%). Het verschil tussen autochtone en allochtone leerlingen is evenwel niet statistisch

significant (chi-kwadraat (2,177) = 4.0, $p > .10$, $cc = 0.24$).

De derde vraag is *waarom* de leerlingen denken dat het discriminerende gedrag dat ze beschrijven zich voordeed.¹ Omdat vanuit de meeste leerlingen discriminatie beschrijven in termen van schelden gaan we alleen in op de resultaten voor dit gedrag. De meerderheid van de leerlingen (58%) geeft geen verklaring voor het discriminerende gedrag, maar beschrijft alleen een voorbeeld van schelden. Van degenen die wel een verklaring geven noemt 12% huidskleur, 73% etnische verschillen en kenmerken (zoals 'buitenlander', 'Turk', 'zwarte' etc.) en 15% persoonlijke en fysieke kenmerken, zoals lichaamsgewicht en uiterlijk. Er zijn geen statistisch significante verschillen tussen autochtone en allochtone leerlingen voor de totale percentages (chi-kwadraat (3,177) = 0.7, $p > .10$, $cc = 0.08$).

3.2 De twaalf verhaaltjes

In Tabel 1 staan de percentages van de leerlingen die bevestigend hebben geantwoord op de vraag of het in de verhaaltjes om discriminatie gaat. Aangezien het gaat om dichotome data (ja of nee discriminatie) hebben we een loglineaire analyse uitgevoerd met etniciteit van de respondent (autochtoon-allochtoon), type verhaaltje (uitschelden, verdelen en uitsluiten) en dader-slachtoffer (de vier configuraties) als factoren. Er is een significant effect voor type

Tabel 1
Percentage leerlingen dat het gedrag in de verhaaltjes beoordeelt als een geval van discriminatie

	TOTALE GROEP (N = 177)	AUTOCHTONEN (N = 103)	ALLOCHTONEN (N = 74)
SCHELDEN			
Autocht-Autocht ¹	83%	84%	83%
Allocht-Allocht	80%	80%	85%
Autocht-Allocht	99%	99%	99%
Allocht-Autocht	83%	89%	78%
VERDELEN			
Autocht-Autocht	34%	34%	30%
Allocht-Allocht	35%	33%	37%
Autocht-Allocht	78%	80%	74%
Allocht-Autocht	50%	50%	51%
UITSLUITEN			
Autocht-Autocht	19%	20%	19%
Allocht-Allocht	23%	24%	20%
Autocht-Allocht	39%	37%	39%
Allocht-Autocht	24%	24%	30%

¹ Het eerste woord verwijst naar de dader in het verhaaltje en het tweede naar het slachtoffer.

verhaaltje (gestandaardiseerde z-waarde is 15.05) en voor vader-slachtoffer configuratie (z-waarde is -5.47). Het effect voor etniciteit is niet significant (z-waarde is 0.70).

In Tabel 1 is te zien dat de verhaaltjes over schelden door de meeste leerlingen als discriminatie worden beoordeeld. Voor de verhaaltjes over het verdelen van snoep en speelgoed zijn de percentages lager en voor sociale uitsluiting zijn ze het laagste. Alle drie de onderlinge verschillen zijn significant (paarsgewijze t-test).

Ten tweede is in Tabel 1 te zien dat binnen ieder type gedrag het verhaaltje met het autochtone kind als vader en het allochtone kind als slachtoffer het hoogste percentage instemmende antwoorden heeft. Hierbij wordt de situatie waarin een autochtoon kind een allochtoon kind uitscheldt door vrijwel alle kinderen als discriminatie beoordeeld.

De loglineaire analyse geeft ook een significant interactie effect aan tussen type verhaaltje en vader-slachtoffer configuratie (z-waarde is 2.03). Bij de verhaaltjes over de verdeling van snoep en speelgoed komt een verschil naar voren: het verhaaltje met een allochtone vader en een autochtoon slachtoffer wordt door een significant hoger percentage als discriminatie beoordeeld in vergelijking met de twee intra-etnische verhaaltjes. Dit verschil komt niet naar voren voor 'uitschelden' en 'uitsluiting'.

3.3 Verklaringen van discriminatie

De leerlingen hebben voor ieder van de twaalf verhaaltjes toegelicht waarom ze het wel of geen voorbeeld van discriminatie vinden. De gecodeerde redenen die zijn gegeven kunnen geanalyseerd worden in termen van leerling- en categorie-frequenties. Aangezien de verklaringen slechts één keer zijn gecodeerd voor iedere leerling komen beide frequenties echter overeen. Daarom zijn categorie-frequenties berekend voor ieder van de twaalf verhaaltjes. Vervolgens zijn percentages berekend voor iedere categorie binnen ieder van de drie typen gedrag. Tabel 2 bevat de resultaten van deze analyses. We zullen eerst de resultaten voor de bevestigende antwoorden bespreken. Ten eerste is in Tabel 2 te zien dat er leerlingen zijn die geen verklaring geven. Op alle twaalf verhaaltjes beperkt ongeveer 20% van de leerlingen zich tot een herhaling van wat in de ver-

haaltjes werd beschreven, bijvoorbeeld door te schrijven dat het discriminatie is omdat de een de ander uitscheldt. Ten tweede is er bij alle drie de typen gedrag een substantieel percentage dat verwijst naar etnische kenmerken (bijv. 'Het is discriminatie want zij is Turks'). Ten derde heeft ieder type gedrag een andere categorie met het hoogste percentage van het totaal aantal gecodeerde antwoorden. Voor 'schelden' is dit de categorie met antwoorden die er de nadruk op leggen dat het om discriminatie gaat omdat het slachtoffer niet verantwoordelijk is. Voor de verhaaltjes over 'verdelen' wordt met name het principe van verdelende rechtvaardigheid gebruikt en voor 'uitsluiten' wordt het principe van gemeenschappelijkheid het meeste genoemd. Er zijn hierbij geen significante verschillen tussen autochtone en allochtone leerlingen. Als we evenwel de percentages voor de autochtone en allochtone leerlingen op ieder verhaaltje apart bekijken komen er twee verschillen naar voren. Een lager percentage van de autochtone leerlingen refereert aan etnische kenmerken bij het verhaaltje waar een autochtoon kind een allochtoon kind uitscheldt (χ^2 -kwadraat (6,177) = 12.2, $p < .05$, $cc = 0.28$). En een hoger percentage van de autochtone leerlingen verwijst naar het principe van verdelende rechtvaardigheid bij het verhaaltje waar een autochtoon kind minder snoep geeft aan een allochtoon kind (χ^2 -kwadraat (6,177) = 16.3, $p < .01$, $cc = 0.35$). De allochtone leerlingen verwezen meer naar etnische kenmerken om dit gedrag te verklaren.

Voor de ontkennende antwoorden is er ook zo'n 20% die geen verklaring geeft voor het antwoord, maar voornamelijk herhaalt wat in het verhaaltje werd beschreven. Ook hier is voor ieder type gedrag een andere categorie met het hoogste percentage van het totaal aantal gecodeerde antwoorden. Voor schelden is dit de categorie verantwoordelijkheid. Als het slachtoffer op een of andere manier zelf verantwoordelijk wordt gehouden dan wordt het gedrag niet als discriminatie gezien. Veronderstelde verantwoordelijkheid van het slachtoffer is ook een rechtvaardiging die bij de andere twee typen gedrag wordt gebruikt. Voor het verdelen van snoep en speelgoed heeft de categorie 'verzachting' het hoogste percentage en deze categorie heeft ook een relatief hoog per-

Tabel 2

Percentages voor iedere categorie binnen de drie typen van gedrag, afzonderlijk voor bevestigende en ontkennende antwoorden op de vraag 'is dit discriminatie?'

	SCHULDEN	'JA' DISCRIMINATIE VERDELEN	UITSLUITEN
Moraliteit	14%	6%	13%
Verdelende rechtvaard.	--	44%	--
Gemeenschappelijkheid	--	--	37%
Gedragsalternatieven	11%	8%	4%
Onverantwoordelijk	33%	2%	4%
Etniciteit	18%	19%	20%
Rest-categorie	10%	3%	3%
Geen verklaring	14%	17%	19%
Totaal aantal gecodeerde antwoorden ¹	596	335	182
		'NEE' DISCRIMINATIE	
Vrijheid	5%	3%	32%
Verantwoordelijkheid	39%	15%	13%
Verzachting	2%	42%	29%
Etniciteit	18%	4%	2%
Rest-categorie	13%	15%	8%
Geen verklaring	23%	22%	16%
Totaal aantal gecodeerde antwoorden	81	330	481

1 Het totaal aantal gecodeerde antwoorden voor iedere antwoordcategorie voor de drie typen gedrag

centage bij de verhaaltjes over sociale uitsluiting. In deze antwoorden wordt ter rechtvaardiging aangevoerd dat bijvoorbeeld de gevolgen van het gedrag niet ernstig zijn, dat de dader niets kwaads bedoelt, of dat het gedrag anders geïnterpreteerd moet worden ('plagen'). De verhaaltjes over sociale uitsluiting worden met name beoordeeld als geen gevallen van discriminatie omdat gesteld wordt dat ieder kind zelf mag beslissen met wie hij of zij wil spelen. Voor de redenen die gegeven worden bij de ontkennende antwoorden komen er geen significante verschillen naar voren tussen autochtonen en allochtonen leerlingen. Tenslotte komt alleen bij de verhaaltjes over schelden een substantieel percentage voor bij de categorie 'etnische kenmerken'. Door deze leerlingen wordt verklaard dat het in het verhaaltje niet om discriminatie gaat omdat de dader en slachtoffer tot dezelfde etnische groep behoren of omdat de dader een allochtoon kind is en het slachtoffer een autochtoon kind.

4 Discussie/conclusies

In dit artikel is ingegaan op de betekenis van het begrip discriminatie voor kinderen zelf. Een

ruime meerderheid van de onderzochte tien- tot en met dertienjarigen weet wat discriminatie is. Bovendien komen er geen leeftijdsverschillen naar voren bij de betekenissen die discriminatie heeft. Leerlingen van tien jaar hebben eenzelfde beeld en idee van discriminatie als leerlingen van dertien jaar. Het prototypische voorbeeld van discriminatie is een situatie waarin een autochtoon kind een allochtoon kind uitscheldt, veelal zonder dat daarvoor een gerechtvaardigde reden bestaat. Op verschillende kenmerken van dit prototype kunnen we nader ingaan. Ten eerste werd in de voorbeelden die de leerlingen zelf gaven discriminatie vooral gezien als gedrag tussen twee of meer kinderen. Discriminatie werd geconceptualiseerd op het interpersoonlijke niveau en niet op het meer abstracte maatschappelijke niveau zoals de arbeidsmarkt of als maatschappelijke ongelijkheid. Ten tweede werd discriminatie voornamelijk gezien als schelden en in mindere mate als oneerlijk delen en sociale uitsluiting (vgl. Troyna & Hatcher, 1992). Discriminatie kan voor de leerlingen naar verschillende vormen van gedrag verwijzen, maar schelden is het meest typische voorbeeld.² Ten derde werd in de eigen voorbeelden van de leerlingen discriminatie voornamelijk gezien als etnische dis-

criminatie en niet bijvoorbeeld als discriminatie op grond van sekse of leeftijd. Deze beperking van discriminatie tot etnische discriminatie is ook bij adolescenten geconstateerd (Leeman, 1994). Ten vierde werd gedrag als discriminerend beoordeeld als het tegen morele beginselen ingaat of als er geen rechtvaardiging is voor het gedrag. Discriminatie is gedrag dat je gewoon niet doet omdat het verkeerd is of omdat het ingaat tegen principes van verdelende rechtvaardigheid en gemeenschappelijkheid. Dit wijst op de morele betekenis van discriminatie en het gebruik van algemene principes bij morele oordelen. Maar gedrag kan ook als discriminerend beoordeeld worden omdat er geen goede reden voor het gedrag is. Er kunnen omstandigheden zijn die het gedrag rechtvaardigen waardoor er geen sprake is van discriminatie. Verschillende leerlingen die bevestigend antwoordden op de vraag of het in het verhaaltje om discriminatie gaat, verklaarden dat de dader gedragsalternatieven had of dat het slachtoffer niet zelf verantwoordelijk is. Hiermee verwezen deze leerlingen impliciet naar mogelijke rechtvaardigingen: als er geen gedragsalternatieven zijn of het slachtoffer is (gedeeltelijk) wel zelf verantwoordelijk dan is er geen sprake van discriminatie. Een dergelijke redenering werd expliciet gevolgd door de leerlingen die het gedrag in de verhaaltjes geen vorm van discriminatie vonden. Zij stelden bijvoorbeeld dat het slachtoffer zelf verantwoordelijk is, of dat de dader het niet meende en dat het allemaal niet zo erg is. Daarnaast verwees een aantal van deze leerlingen naar het principe van vrijheid om het gedrag te rechtvaardigen. Dit wijst er op dat abstracte principes ook kunnen worden gebruikt om bijvoorbeeld uitsluiting te rechtvaardigen en niet alleen om gedrag te veroordelen.

Discriminatie werd overwegend gezien als een onaanvaardbare handeling van een autochtoon kind ten opzichte van een allochtoon kind. Voor sommige leerlingen was dit zelfs een definiërend kenmerk van discriminatie omdat ze hetzelfde gedrag tussen kinderen van een zelfde etnische groep of een allochtone dader en een autochtoon slachtoffer, niet als discriminatie beoordeelden. Deze resultaten laten zien dat de beoordeling van discriminatie niet alleen een kwestie is van de toepassing van abstracte morele principes. Kennis in de vorm van gang-

bare voorstellingen zijn van belang. Discriminatie is iets wat autochtone kinderen allochtone kinderen aandoen en met name uitschelden zonder reden. In termen van sociale representaties is dit de kern van de voorstelling die kinderen van discriminatie hebben. Een dergelijke kern wordt verondersteld de inhoud van de representatie te organiseren en de betekenis ervan te genereren (Guimelli, 1993). Door de organiserende en betekenisgevende functies komt deze kern over als onweerlegbaar en vanzelfsprekend waardoor alle leerlingen er mee instemmen dat het om discriminatie gaat.

In ons onderzoek kwamen er weinig verschillen naar voren tussen autochtone en allochtone leerlingen in hun begrip van discriminatie. Dit is tegengesteld aan het idee dat allochtone kinderen een ander begrip van discriminatie hebben op grond van de maatschappelijke positie van hun groep en eventuele ervaringen met negatieve oordelen over hun groep of henzelf. Maar het stemt overeen met het idee van sociale representaties dat benadrukt dat kinderen uit een zelfde sociale omgeving gemeenschappelijke voorstellingen hebben die ze gebruiken om (elkaars) gedrag van betekenis te voorzien. Er komen evenwel in ons onderzoek ook enkele significante verschillen tussen autochtone en allochtone leerlingen naar voren. Zo beschreven allochtone leerlingen eerder een allochtoon kind als slachtoffer en verwezen ze meer naar etnische kenmerken als verklaring voor discriminerend gedrag. Dit geeft aan dat het behoren tot een sociale meerderheids- of minderheidsgroepering niet geheel zonder belang is voor het begrip dat kinderen hebben van discriminatie.

Ondanks deze verschillen springen de vele overeenkomsten in het oog. Wij hebben geprobeerd te laten zien dat deze overeenkomsten geïnterpreteerd kunnen worden in termen van sociale representaties van discriminatie. Het lijkt bij de betekenis van discriminatie te gaan om sociaal gedeelde opvattingen, ideeën en waarden die georganiseerd zijn rond enkele betekenisdragende elementen en die worden ge(re)produceerd binnen de gemeenschap van kinderen en hun peergroep (Corsaro & Eder, 1990). Deze interpretatie betekent natuurlijk niet dat leerprocessen of algemene ontwikkelingen in het denkvermogen van kinderen van

geen belang zijn. Integendeel, maar cognitieve ontwikkelingsmodellen geven weinig inzicht in de inhoud van het denken van kinderen. De nadruk ligt op processen en structuren en niet zozeer op datgene waarover kinderen denken en hoe ze daarbij gangbare ideeën en opvattingen gebruiken.

Tenslotte zijn twee opmerkingen van belang bij het idee van sociale representaties en het uitgevoerde onderzoek. Beide opmerkingen verwijzen naar beperkingen van ons onderzoek en daarmee naar mogelijkheden voor vervolgonderzoek. Ten eerste zijn we er vanuit gegaan dat kinderen die in een zelfde sociale omgeving leven gemeenschappelijke ideeën en voorstellingen hebben. De sociale omgeving zelf is echter niet onderzocht en dit uitgangspunt is ook niet noodzakelijk juist. Hewitt (1986) laat bijvoorbeeld zien dat opgroeiende kinderen in eenzelfde buurt niet vanzelfsprekend leidt tot gedeelde betekenisverlening aan discriminatie. In de twee buurten die hij onderzoekt is dat in het ene geval wel en in het andere geval niet zo (zie ook Back, 1993). De mate waarin en de wijze waarop kinderen daadwerkelijk met elkaar omgaan zodat gemeenschappelijke betekenissen ge(re)produceerd kunnen worden verdient dan ook nader aandacht.

Ten tweede richt het idee van sociale representaties zich vooral op het verklaren van overeenkomsten in het denken. Dit brengt met zich mee dat er gemakkelijk onvoldoende oog is voor individuele verschillen. Niet alle kinderen in ons onderzoek zijn het over alle aspecten eens. Als constructivistische benaderingen het gevaar lopen van een eenzijdig individualistisch perspectief op cognitieve processen dan lopen theorieën over sociale cognitie het gevaar van een eenzijdig collectief perspectief. Onderzoek naar sociale representaties dient met andere woorden meer aandacht te krijgen voor de relatie tussen gedeelde betekenissen en individuele variatie (zie De Rosa, 1994).

Afsluitend hebben wij een nauwkeurig beeld proberen te geven over wat kinderen onder discriminatie verstaan. Dit roept niet alleen allerlei interessante theoretische vragen op, maar kan wellicht ook helpen om in de onderwijspraktijk beter aan te sluiten bij de beleavingswereld van kinderen.

Noten

- 1 Een vierde vraag is waar en wanneer het gedrag zich voordoet. Omdat de meeste kinderen hierover in hun beschrijving geen informatie hebben gegeven kunnen we hier niet nader op ingaan.
- 2 In tegenstelling tot hun leerlingen zagen enkele leerkrachten van de scholen waar het onderzoek is uitgevoerd uitschelden niet als discriminatie. Zij vertelden ons vooraf dat er maar weinig gevallen van discriminatie tussen kinderen voorkomen, maar dat er wel regelmatig gescholden werd.

Literatuur

- Aboud, F. (1988). *Children and prejudice*. New York: Basil Blackwell.
- Back, L. (1993). Race, identity and nation within an adolescent community in south London. *New Community*, 19, 217-233.
- Clark, K. (1966). *Prejudice and your child* (2nd ed.) Boston: Beacon.
- Corsaro, W. (1992). Interpretive reproduction in children's peer cultures. *Social Psychology Quarterly*, 55, 160-177.
- Corsaro, W. A., & Eder, D. (1990). Children's peer culture. *Annual Review of Sociology*, 16, 197-220.
- Damon, W. (1977). *The social world of the child*. San Francisco: Jossey-Bass.
- Duveen, G., & Lloyd, B. (Eds.) (1990). *Social representations and the development of knowledge*. Cambridge: Cambridge University Press.
- Elmer, N. (1987). Socio-moral development from the perspective of social representations. *Journal for the Theory of Social Behaviour*, 17, 371-388.
- Emler, N., & Dickinson, J. (1985). Children's representation of economic inequalities: The effects of social class. *British Journal of Developmental Psychology*, 3, 191-198.
- Emler, N., & Ohana, J. (1993). Studying social representations in children. In G. M. Breakwell & D. V. Canter (Eds.), *Empirical approaches to social representations* (pp. 63-89). Oxford: Clarendon.
- Emler, N., Ohana, J., & Moscovici, S. (1987). Children's beliefs about institutional roles: A cross-national study of representations of the teacher's role. *British Journal of Educational Psychology*, 57, 26-37.
- Guimelli, C. (1993). Locating the central core of social representations: towards a method. *European Journal of Social Psychology*, 23, 555-559.

Hewitt, R. (1986). *White talk black talk*. Cambridge: University Press.

Leeman, Y. (1994). *Samen jong: Nederlandse jongeren en lessen over inter-etnisch samenleven en discriminatie*. Utrecht: Van Arkel.

Lloyd, B. (1987). Social representations of gender. In J. Bruner & H. Haste (Eds.), *Making sense: The child's construction of the world* (pp. 147-162). London: Methuen.

Milner, D. (1983). *Cildren and race: Ten years on* (2nd ed.) London: Ward Lock.

Moscovici, S. (1982). The coming era of social representations. In J. P. Codol & J. Leyens (Eds.), *Cognitive analysis of social behavior* (pp. 115-150). The Hague: Nijhoff.

Moscovici, S. (1984). The phenomenon of social representations. In R. Farr & S. Moscovici (Eds.), *Social representations* (pp. 3-69). Cambridge: Cambridge University Press.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.) (1991). *Perspectives on socially shared cognition*. Washington: American Psychological Association.

Rosa, A. S. De (1994). From theory to metatheory in social representations: The lines of argument of a theoretical-methodological debate. *Social Science Information*, 33, 273-304.

Saharso, S. (1992). *Jan en alleman: Etnische jeugd over etnische identiteit, discriminatie en vriendschap*. Utrecht: Van Arkel.

Troyna, B., & Hatcher, R. (1992). *Racism in children's lives: A study of mainly-white primary schools*. London: Routledge.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.

Uzzell, D., & Blut, L. (1993). Vikings! Children's social representations of history. In G. M. Breakwell & D. C. Canter (Eds.), *Empirical approaches to social representations* (pp. 110-133). Oxford: Clarendon.

Verkuyten, M. (1992). *Zelfbeleving van jeugdige allochtonen*. Lisse: Swets & Zeitlinger.

Verkuyten, M., & Masson, K. (1993). Etnisch categoriseren door kinderen op multi-etnische scholen: het relatieve belang van etnische achtergrond. *Psychologie & Maatschappij*, 63, 147-161.

Zeegers, W., & Jansz, J. (1988). Betekenisgeving als sociaal proces. *Psychologie & Maatschappij*, 12, 117-132.

Auteurs

M. Verkuyten, B. Kinket en Ch. van der Wielen zijn verbonden aan de vakgroep Algemene Sociale Wetenschappen van de Sociale Faculteit aan de Universiteit Utrecht.

Adres: Heidelberglaan 2, 3584 CS Utrecht.

Abstract

Children's understanding of discrimination

M. Verkuyten, B. Kinket & Ch. van der Wielen. *Pedagogische Studiën*, 1996, 73, 447-458.

In total 193 Dutch and ethnic minority children between 10 and 13 years of age participated in this study. Using a questionnaire with open-ended and closed questions it was found that there were shared beliefs and understandings about when a specific act is considered discriminatory, who the actors are and why the discrimination occurs. The prototypical example of discrimination is a situation in which a Dutch child calls a minority child names without an acceptable reason. Discrimination was seen to a lesser degree as unequal division of valued things and as social exclusion. Consistent with social representation theory and contrary to common belief there were very few differences in understanding of discrimination between ethnic majority and minority children. No differences were found for age and gender.

Leereffecten van Jeugdjournaal en gedrukt nieuws bij kinderen: Een mediumvergelijkend experiment*

J. H. Walma van der Molen en T. H. A. van der Voort

Samenvatting

In deze studie werd een experiment uitgevoerd waarin (a) de herinnering van kinderen van nieuws gepresenteerd via de televisie en in gedrukte vorm werd vergeleken en (b) werd nagegaan of de relatieve effectiviteit van beide media als overbrengers van nieuws afhankelijk is van het leesvaardigheidsniveau van kinderen en een tevoren gecreëerde toetsverwachting. Aan een steekproef van 152 kinderen uit groep 6 en 8 werd een compilatie van vijf onderwerpen uit het Jeugdjournaal voorgelegd, ofwel in de vorm van de originele televisieberichten, ofwel in de vorm van inhoudelijk vergelijkbare gedrukte versies. Binnen iedere conditie wist de helft van de kinderen dat een retentietoets zou volgen; de andere kinderen hadden geen toetsverwachting. De resultaten op een schriftelijke kennistoets lieten zien dat de kinderen die het nieuws op de televisie zagen zich meer herinnerden dan de kinderen die de gedrukte versies lazen, ongeacht hun leesvaardigheidsniveau of toetsverwachting.

1 Inleiding

De afgelopen decennia is de televisie voor de meeste mensen de belangrijkste bron voor nieuwsinformatie geworden. Al vanaf het midden van de jaren zestig werd deze trend zichtbaar in publieke-opinie-studies en gedurende de laatste dertig jaar heeft de televisie haar voorsprong op haar naaste mededinger, de krant, geleidelijk aan steeds verder vergroot (Roper Reports, 1995). Het feit dat veel mensen zeggen dat de televisie hun primaire

nieuwsbron is, betekent echter nog niet dat de televisie ook daadwerkelijk het meeste bijdraagt aan de kennis die het publiek van het nieuws heeft. In de eerste plaats zijn er aanwijzingen dat respondenten de mate waarin zij aan televisienieuws en aan kranten worden blootgesteld soms verkeerd inschatten (Robinson & Levy, 1986). Bovendien twijfelen veel onderzoekers aan het vermogen van de televisie om nieuwsinformatie over te dragen en wordt de krant meestal als een effectiever medium gezien (e.g., Gunter, 1987; Robinson & Levy, 1986).

Drie soorten redenen zijn aangevoerd waarom men van televisienieuws minder kennis zou opdoen dan van nieuws uit de krant. De eerste soort heeft betrekking op *intrinsieke verschillen* tussen de twee media en de wijze waarop de verstrekte informatie cognitief wordt verwerkt. Eén intrinsiek verschil is dat kijkers over hun informatieverwerking minder controle kunnen uitoefenen dan lezers (Kozma, 1991). Terwijl het lezen van gedrukt nieuws in een zelf-gekozen tempo plaatsvindt, wordt het tempo waarin men het televisienieuws consumeert bepaald door de programmamaker. Vanaf het moment dat een onderwerp is afgelopen, heeft de kijker niet meer de gelegenheid om terug te kijken. De lezer heeft daarentegen de kans om passages te herlezen en details na te kijken, hetgeen het opslaan van informatie kan bevorderen (e.g., Furnham & Gunter, 1985; Gunter, 1987). Een tweede intrinsiek verschil betreft de wijze waarop beide media het nieuws structureren. De opbouw van kranteberichten is beter zichtbaar dan die van nieuwsberichten op de televisie. Het gebruik van kopjes, de indeling in paragrafen en andere middelen die bij gedrukt materiaal gebruikt worden om lijn in een verhaal aan te brengen, kunnen het verwerken van informatie uit de krant bevorderen (Furnham, Gunter & Green, 1990). Tenslotte presenteert de televisie het nieuws met behulp

* Dit onderzoek werd gesubsidieerd door de Stichting Gedragswetenschappen (SGW) van NWO (projectnummer 575-87-706).

van zowel bewegende beelden als geluid en gesproken commentaar, terwijl de krant zich beperkt tot het gedrukte woord, eventueel aangevuld met fotomateriaal. In beginsel kunnen de beelden die de televisie aan de verbale boodschap toevoegt de overdracht van informatie bevorderen (Gunter, 1991). Dit voordeel van de televisie gaat echter alleen op wanneer de beelden inhoudelijk sporen met het gesproken commentaar.

De tweede soort verklaringen heeft betrekking op de manier waarop *producenten* van beide media gebruik maken. Het voordeel dat het gebruik van televisiebeelden biedt, wordt vaak weer ongedaan gemaakt door de wijze waarop de makers van televisienieuws het medium gebruiken. De in het televisienieuws vertoonde beelden zijn vaak ongerelateerd aan het gesproken commentaar, hetgeen de kijker juist van de verbale boodschap kan afleiden (Drew & Grimes, 1987; Grimes, 1991; Reese, 1984; Wember, 1976). Daarnaast wordt het televisienieuws vaak in een hoog tempo aangeboden, met weinig redundantie en herhaling, waardoor de kijker moeite heeft de informatie tijdig in zich op te nemen (e.g. Robinson & Levy, 1986). Tenslotte ontvangt de kijker per bericht beduidend minder informatie dan de lezer van de krant (Robinson & Levy, 1986). In nieuwsuitzendingen wordt meestal niet meer dan twee minuten voor een onderwerp ingeruimd (Grabber, 1990) en een complete journaaluitzending van 25 tot 30 minuten bevat minder verbale informatie dan de voorpagina van een dagblad (Furnham & Gunter, 1985).

Een derde type verklaring wijt het veronderstelde geringere vermogen van de televisie om nieuwsinformatie over te dragen aan verschillen in de wijze waarop de *consumenten* van beide media gebruik maken. Vaak wordt verondersteld dat de kijker het nieuws in vergelijking met de krantelezer minder aandachtig tot zich neemt en gemakkelijker wordt afgeleid (Robinson & Levy, 1986).

De aanname dat de kijker naar het televisienieuws minder kennis opdoet dan de lezer van de krant, wordt ondersteund door correlatieve studies waarin de gebruiksfrequentie van de twee nieuwsmedia gerelateerd werd aan de mate waarin het publiek op de hoogte was met het nieuws. Krantelers kwamen uit de meeste van deze studies als beter geïnformeerd naar

voren dan kijkers, een verschil dat overeind bleef na controle voor opleidingsniveau en andere predictoren (zie voor een overzicht: Robinson & Davis, 1990; Robinson & Levy, 1986). Dit verschil in informatieniveau tussen kijkers en lezers is echter deels verklaarbaar uit het feit dat de krant doorgaans meer informatie over het nieuws geeft dan de televisie. Van dit voordeel was echter geen sprake in experimenten waarin een vergelijking werd gemaakt tussen het onthouden van nieuwsberichten van de televisie en de herinnering van gedrukte versies die identiek waren aan het gesproken televisiecommentaar. Met uitzondering van één studie, waarin geen mediumverschil werd aangetroffen (Stauffer, Frost & Rybolt, 1981), lieten alle mediavergelijkende experimenten zien dat de gedrukte versies beter werden onthouden dan de vergelijkbare televisieversies (DeFleur, Davenport, Cronin & DeFleur, 1992; Facorro & DeFleur, 1993; Furnham & Gunter, 1985; Gunter & Furnham, 1986; Gunter, Furnham & Gietson, 1984; Gunter, Furnham & Leese, 1986; Wicks & Drew, 1991; Wilson, 1974).

Deze experimentele studies naar de relatieve verdiensten van televisie en krant als overdragers van nieuws zijn nagenoeg uitsluitend uitgevoerd bij volwassenen. Afgezien van één quasi-experimentele mediumvergelijkende studie (Beentjes, Vooijs & Van der Voort, 1993), is tot op heden niet nagegaan of de gevonden superioriteit van de krant ook voor kinderen geldt. Anders dan bij volwassenen, is bij kinderen in ieder geval wel duidelijk dat de televisie inderdaad de hoofdbron is voor nieuwsinformatie. Ongeveer tweederde van de kinderen uit de hoogste klassen van de basisschool zegt dagelijks of wekelijks naar nieuwsprogramma's op de televisie te kijken (Anderson, Mead & Sullivan, 1986). De overgrote meerderheid van de kinderen noemt de televisie als het medium waaraan zij de meeste kennis omtrent het nieuws ontleen, aanzienlijk meer dan aan andere media en aan discussies met leraren, leeftijdgenoten of ouders (Comstock & Paik, 1991). Aan de juistheid van deze opinie van kinderen hoeft niet getwijfeld te worden, omdat onderzoek heeft aangetoond dat kinderen nauwelijks kranten lezen, de radio primair als 'music box' gebruiken (Van Lil, 1989) en zelden met anderen over het nieuws praten (Drew & Reeves, 1980). Omdat kinderen nauwelijks

alternatieve nieuwsbronnen gebruiken, heeft de komst van de televisie de toegankelijkheid van nieuws voor kinderen sterk vergroot (Comstock & Paik, 1991).

De eerder besproken mediumvergelijkende experimenten met volwassenen suggereren dat het nadelig is dat kinderen vooral via de televisie met het nieuws geconfronteerd worden, omdat de krant effectiever zou zijn bij de overdracht van nieuwsinformatie. Het is echter zeer de vraag of de bevindingen uit de bij volwassenen uitgevoerde mediavergelijkende experimenten op kinderen van toepassing zijn. Er zijn tenminste drie factoren die de generaliseerbaarheid van de resultaten van deze experimenten beperken. In de eerste plaats is de leesvaardigheid bij kinderen minder sterk ontwikkeld dan bij volwassenen, waardoor kinderen waarschijnlijk minder goed in staat zijn om te profiteren van de voordelen die gedrukt nieuws bij de opname van informatie biedt. Daarnaast zijn de mediumvergelijkende studies uitsluitend uitgevoerd bij hoger opgeleide volwassenen. Dit roept de vraag op of de veronderstelde superioriteit van de krant niet beperkt blijft tot of vooral opgaat voor de goede lezers, terwijl de matige lezers relatief meer baat zouden kunnen hebben bij nieuws dat verstrekt wordt via de televisie. Tenslotte zijn de mediumvergelijkende studies uitgevoerd bij proefpersonen die te voren wisten dat zij na afloop over hun kennis van het nieuws zouden worden ondervraagd. De bevindingen uit deze studies zijn daarom niet noodzakelijkerwijze van toepassing op de thuissituatie, waar beide media gebruikt worden zonder dat een toets wordt verwacht. Denkbaar is, dat de aankondiging van een retentietoets verschillend uitwerkt op de motivatie van lezers en kijkers, hetgeen implicaties kan hebben voor de relatieve effectiviteit die voor beide media wordt aangetroffen.

De bedoeling van de huidige studie was driedelig. Het eerste, en meest belangrijke doel, was na te gaan of het bij hoog opgeleide volwassenen gevonden leervoordeel van gedrukt nieuws boven televisienieuws ook geldt voor kinderen uit de hoogste klassen van de basisschool. Meer in het bijzonder werd onderzocht hoe de herinnering van kinderen van televisienieuws dat speciaal voor hen is gemaakt (het Jeugdjournaal) zich verhoudt tot de retentie van inhoudelijk vergelijkbaar gedrukt

nieuws. Ten tweede werd onderzocht in hoeverre de relatieve effectiviteit van televisie en krant als overdragers van nieuws afhankelijk is van het niveau van leesvaardigheid van de kinderen. Ten slotte werd nagegaan of de relatieve effectiviteit van televisie- en gedrukt nieuws afhankelijk is van de wetenschap dat men over de opgedane kennis van het nieuws ondervraagd zal worden.

1.1 Hypothesen

Differentiële herinnering van televisie- en verbaal nieuws

Hoe de differentiële herinnering van televisie- en gedrukt nieuws bij kinderen uitvalt, is tot op heden slechts in één quasi-experimenteel mediumvergelijkend experiment onderzocht (Beentjes, Vooijs & Van der Voort, 1993). Beentjes et al. vonden bij kinderen uit de hoogste klassen van de basisschool dat kijkers naar een compilatie van onderwerpen uit het Jeugdjournaal zich meer herinnerden dan lezers van inhoudelijk identieke gedrukte versies, zij het dat het aangetroffen leervoordeel van televisie beperkt bleef tot de kinderen die geen retentietoets verwachtten.

In enkele mediumvergelijkende experimenten die eveneens bij kinderen uit de hoogste klassen van het basisonderwijs werden uitgevoerd, is nagegaan hoe de herinnering van *fictieve* televisieverhalen zich verhoudt tot de herinnering van gedrukte versies. Deze studies lieten eveneens zien dat kinderen niet noodzakelijkerwijs het meest gebaat zijn bij gedrukte informatie. Terwijl in de Israëlische studies van Salomon en Leigh (1984) geen significante mediumverschillen in herinnering werden aangetroffen, wees een tweetal Nederlandse studies uit dat kinderen fictieve televisieverhalen, vergeleken met gedrukte versies, met minder fouten (Beentjes & Van der Voort, 1991a; 1991b) en vollediger (Beentjes & Van der Voort, 1991b) navertelden. In een derde Nederlandse studie was het leervoordeel van televisie boven tekst echter alleen waarneembaar op een uitgestelde herinneringstoets (Beentjes & Van der Voort, 1993).

Naar het onthouden door kinderen van fictieve verhalen zijn behalve studies waarbij televisie- en gedrukt materiaal werd vergeleken ook studies uitgevoerd waarbij een vergelijking tussen televisie en radio werd gemaakt. In

deze televisie-radio-studies werd de herinnering van fictieve televisieverhalen geconstrueerd met de retentie van inhoudelijk identieke verhalen die kinderen beluisterden. De bevindingen uit deze studies kunnen indicatief zijn voor de te verwachten uitkomsten bij de huidige vergelijking van televisie- en gedrukt nieuws, omdat er aanwijzingen zijn dat de vaardigheden die kinderen nodig hebben om beluisterde verhalen te verwerken overeenkomstig zijn aan de vaardigheden benodigd voor de verwerking van zelf-gelezen verhalen (Pezdek, Lehrer & Simon, 1984). Pezdek et al. boden kinderen een fictief verhaal in gedrukte vorm aan en een ander verhaal in ofwel een radio- ofwel een televisieuitvoering. De retentie van de gedrukte verhalen vertoonde een positieve samenhang met die van de beluisterde verhalen, maar was ongerelateerd aan de retentie van de televisieverhalen. De auteurs leidden hieruit af dat er een overlap bestaat tussen de informatieverwerkingsvaardigheden waarop de twee verbale media, radio en het gedrukte woord, een beroep doen.

De televisie-radio-studies lieten bijna zonder uitzondering zien dat kinderen het meest onthielden van de televisieversies: de kinderen uit de televisieconditie herinnerden zich meer centrale informatie (Greenfield & Beagles-Roos, 1988), meer verhaaldetails (Beagles-Roos & Gat, 1983; Greenfield & Beagles-Roos, 1988) en meer acties (Beagles-Roos & Gat, 1983; Gibbons, Anderson, Smith, Field & Fischer, 1986; Hayes, Kelly & Mandel, 1986; Meringoff, 1980). De bevinding dat televisieverhalen beter werden onthouden dan radioverhalen is vaak verklaard met behulp van de duale-coderings-hypothese van Paivio (1969, 1971). Volgens Paivio wordt audiovisuele informatie in twee verschillende maar gerelateerde codes in het geheugen opgeslagen, een verbale en een visuele code, terwijl radioinformatie alleen verbaal wordt opgeslagen. Tijdens de herinnering kan de visuele code als een extra herinneringscode dienen bovenop de verbale code en daarom zouden televisieverhalen beter worden onthouden dan radioverhalen.

In de huidige studie werd eveneens verwacht dat de televisie het meest effectieve medium is om aan kinderen informatie over te dragen, en wel om drie redenen. Ten eerste, omdat, zoals eerder gesteld, kinderen door hun nog

niet volledig ontwikkelde leesvaardigheid niet in staat zijn om optimaal te profiteren van de mogelijkheden die gedrukt materiaal biedt bij de verwerking van informatie. Ten tweede, omdat de televisie-radio-studies hebben aangetoond dat kinderen juist wel baat hebben bij de extra geheugensteun die de toevoeging van beelden aan het gesproken commentaar bij de televisie biedt. En ten derde, omdat in de huidige studie gebruik werd gemaakt van berichten uit het Jeugdjournaal, een programma waarin ernaar gestreefd wordt om geen beeldmateriaal te gebruiken dat afleidt van de verbale boodschap. De eerste hypothese luidde daarom:

H₁: Kinderen onthouden meer van televisienieuws dan van vergelijkbare gedrukte nieuwsversies.

Leesvaardigheid

Goede lezers onderscheiden zich van matige lezers doordat zij beter in staat zijn om zowel centrale als bijkomstige informatie uit een tekst op te nemen, een tekst te analyseren en er gepaste conclusies uit te trekken (Daneman, 1991). Omdat goede lezers teksten beter onthouden dan matige lezers, ligt het in de rede dat het door ons verwachte verschil in de herinnering van televisie- en gedrukt nieuws bij goede lezers minder sterk ten gunste van televisie zal uitvallen dan bij matige lezers. Bij ons weten is een dergelijk interactie-effect tot op heden slechts in één studie nagegaan (Beentjes & Van der Voort, 1993). De auteurs vonden inderdaad aanwijzingen dat het leervoordeel van televisie boven tekst bij de goede lezers minder sterk was dan bij de matige lezers. Deze bevinding had echter betrekking op het onthouden van fictieve verhalen. Bovendien vond de classificatie van kinderen in goede en matige lezers plaats op basis van leerkrachtoordelen in plaats van objectieve testgegevens. Daarom werd dit interactie-effect in de huidige studie opnieuw onderzocht. De verwachting luidde:

H₂: Het leervoordeel van televisie- boven gedrukt nieuws valt bij matige lezers hoger uit dan bij goede lezers.

In de huidige studie waren twee leeftijdsgroepen betrokken: kinderen uit groep 6 en groep 8 van het basisonderwijs. Omdat de leesvaardigheid van de jongste leeftijdsgroep minder ontwikkeld is die dan van de oudere leef-

tijdsgroep, wordt het gedrukte nieuws door de jongere kinderen vermoedelijk minder goed onthouden dan door de oudere kinderen, met als mogelijke implicatie dat de jongere kinderen relatief meer baat hebben bij de televisie-informatie. Een dergelijk interactie-effect werd inderdaad in een vergelijkingsstudie van Beentjes en Van der Voort (1991b) aangetroffen. Zij vonden dat kinderen uit groep 6 een fictief televisieverhaal beter reproduceerden dan een vergelijkbaar gedrukt verhaal, terwijl het voordeel van televisie boven tekst bij de kinderen uit groep 8 was verdwenen. Derhalve luidde de verwachting:

- H₃: Het leervoordeel van televisie- boven gedrukt nieuws valt bij kinderen uit groep 6 hoger uit dan bij kinderen uit groep 8.

Toetsverwachting

Terwijl in het dagelijks leven de kijker en de krantelezer het nieuws tot zich nemen zonder te verwachten dat zij achteraf over de opgedane kennis ondervraagd worden, wisten de proefpersonen in de mediumvergelijkingsexperimenten met volwassenen steeds op voorhand dat hun retentie van het nieuws getoetst zou worden. Verwacht mag worden dat proefpersonen gemotiveerd door de aankondiging van een toets meer hun best zullen doen om zowel het televisie- als het gedrukte nieuws in zich op te nemen, met als gevolg dat de retentie wordt verhoogd (Craik, 1979). De vraag is echter of de aankondiging van een test de mentale inspanning van kijkers en lezers in dezelfde mate verhoogt. In een studie bij kinderen vonden Beentjes et al. (1993) dat de aankondiging van een retentietoets het mentale inspanningsniveau bij lezers sterker verhoogde dan bij kijkers, een resultaat dat de auteurs niet hadden verwacht. Salomon en Leigh (1984) vonden daarentegen, eveneens bij kinderen, dat het mentale inspanningsniveau van de kijkers door een toetsaankondiging sterker werd verhoogd dan dat van de lezers. Deze uitslag was wel voorzien en de auteurs gaven er ook een plausibele verklaring voor. Omdat kinderen van nature, dat wil zeggen in een situatie zonder toetsverwachting, al geneigd zijn om meer mentale inspanning te investeren in het lezen van teksten dan in het kijken naar de televisie (Salomon, 1984; Salomon & Leigh, 1984), lijkt het

aannemelijk dat de aankondiging van een toets vooral de mentale inspanning van de kijkers zal verhogen. In de huidige studie luidde de verwachting daarom:

- H₄: De hoeveelheid mentale inspanning die kinderen zonder toetsverwachting in het verwerken van televisie- en gedrukt nieuws investeren, wordt door de aankondiging van een toets bij de kijkers sterker verhoogd dan bij de lezers.

Indien vooral de kijkers baat hebben bij een motivatieverhogende instructie, mag tevens verondersteld worden dat de kijkers, aangespoord door een toetsaankondiging, ook meer leren dan de lezers. In de huidige studie werd daarom verwacht:

- H₅: Het leervoordeel van televisie- boven gedrukt nieuws valt bij kinderen die een toets verwachten hoger uit dan bij kinderen die geen toets verwachten.

2 Methode

2.1 Proefpersonen

Het onderzoek werd uitgevoerd bij 73 meisjes en 79 jongens, afkomstig uit de groepen 6 ($n = 72$; M leeftijd = 9.3 jaar) en 8 ($n = 80$; M leeftijd = 11.2 jaar) van vijf basisscholen uit Leiden en omgeving. De kinderen kwamen zowel uit lagere als gemiddelde sociaal-economische milieus. Zij hadden geen leer-, visuele, spraak- of gehoorstoornissen. Kinderen met een score op een gestandaardiseerde test voor begrijpend lezen (CITO, 1981, 1991) die meer dan twee standaardafwijkingen onder het groeps-gemiddelde lag, werden buiten de steekproef gehouden. De aanname was dat deelname van deze subgroep van zeer zwakke lezers de uitkomsten voor de tekstconditie te zeer zou benadelen.

2.2 Opzet

Gebruik werd gemaakt van een 2 (groep 6 vs. groep 8) $\times$ 2 (matige vs. goede lezers) $\times$ 2 (tekst vs. televisie) $\times$ 2 (met vs. zonder toetsverwachting)-opzet. Ter verkleining van de foutenvariantie werd gekozen voor een 'randomized block factorial design' (Kirk, 1968). Begrijpend lezen werd als matchingvariabele gebruikt bij het vormen van blokken, in de verwachting dat deze variabele positief correleert met retentie van het nieuws. Binnen elk van de

twee leeftijdsgroepen werden kinderen aan de hand van hun scores op een test voor begrijpend lezen gegroepeerd in blokken van vier kinderen met identieke of aangrenzende scores. Vervolgens werden de kinderen binnen elk blok willekeurig toegewezen aan één van de 2 (tekst vs. televisie) x 2 (met vs. zonder toetsverwachting) experimentele condities. Omdat de kinderen gegroepeerd waren in blokken van vier met een vergelijkbaar leesvaardigheidsniveau, konden de factoren 'medium' en 'toetsverwachting' in de analyse behandeld worden als binnen-proefpersonen-, of liever, 'binnen-blokken'-factoren; de twee overige factoren waren tussen-proefpersonen-factoren. Het onderscheid tussen matige en goede lezers werd tot stand gebracht door binnen iedere leeftijdsgroep de mediaan bij de scores op de begrijpend lezen test als 'cut-off score' te gebruiken. De helft van de kinderen in het experiment behoorde tot de categorie 'matige lezers', de andere helft tot de 'goede lezers'. Om te controleren voor mogelijke proefleidereffecten, werd binnen elke leeftijdsgroep de deelname van drie proefleiders systematisch gevarieerd over de vier experimentele condities. De belangrijkste afhankelijke variabelen waren: de hoeveelheid mentale inspanning die kinderen in het lezen of bekijken van het nieuws investeren en de proportie goede antwoorden op een kennistoets over de nieuwsberichten.

2.3 Stimulusmateriaal

Elk kind nam kennis van vijf nieuwsberichten. De helft van de kinderen zag de berichten op de televisie; de andere helft las zelf de gedrukte versies van deze berichten.

De vijf onderwerpen waren ontleend aan het Jeugdjournaal, een nieuwsrubriek die het nieuws begrijpelijk wil maken voor kinderen tussen de 10 en 12 jaar. Om voorkennis over de gebruikte nieuwsonderwerpen te verkleinen, werden onderwerpen geselecteerd die ten minste een jaar voor aanvang van het experiment waren uitgezonden. Met hetzelfde doel werden onderwerpen gekozen die slechts één keer in het nieuws waren geweest. Het gesproken commentaar bij de televisie-nieuwsberichten was zonder kennis van de begeleidende beelden begrijpelijk. Daarom kon het letterlijk uitgeschreven commentaar van de televisieversies gebruikt worden in de tekstconditie,

zonder iets weg te laten of toe te voegen.

De vijf geselecteerde berichten vormden een representatieve afspiegeling van het nieuws dat in het Jeugdjournaal aan bod komt, en bestonden dan ook uit zowel berichten die vooral voor kinderen interessant zijn als typische 'volwassen' nieuwsberichten. Zij handelden over de volgende gebeurtenissen: 'Beer vrijgelaten in natuurpark na jaren van gevangenschap', 'Japanse kinderen spijbelen veel omdat ze gepest worden', 'Koningin Beatrix opent nieuwe treintunnel', 'Neushoorn ten onrechte verkocht aan een circus' en 'Nieuwe Amerikaanse wet maakt het moeilijker om wapens te kopen'. Ten behoeve van de televisieconditie werden de vijf onderwerpen aan elkaar gemonteerd en gegoten in het 'format' van een complete uitzending van het Jeugdjournaal, inclusief leader en aftiteling. De totale duur van de televisieversie was 11 minuten, ongeveer gelijk aan de duur van een normale uitzending van het Jeugdjournaal.

De tekstversies van de vijf berichten werden opgemaakt als een artikel uit de krant. De titelkaarten die in de televisieversies werden gebruikt bij de aankondiging van elk nieuw onderwerp werden als kop voor de tekstversies gebruikt. De introductie bij ieder onderwerp, die aan het begin van elk nieuwsbericht door de nieuwslezer werd uitgesproken, werd in de eerste alinea van de tekstversies vet gedrukt weergegeven. De resterende tekst werd afgedrukt in twee kolommen. Op deze manier werd geprobeerd de normale opmaak van een krantartikel zo dicht mogelijk te benaderen, terwijl de verbale informatie in de tekst- en televisieversies constant werd gehouden.

2.4 Procedure

Per deelnemende school werd het experiment tegelijkertijd in twee ongebruikte kamers of lokalen in het schoolgebouw uitgevoerd, een kamer voor de televisieconditie en een kamer voor de tekstconditie. De kinderen werden getest in kleine groepjes van drie of vier leerlingen. De kinderen werden in plaats van ieder apart, in kleine groepjes getest om de vergelijkbaarheid tussen de experimentele conditie zonder toetsverwachting en de thuissituatie te vergroten. De veronderstelling was dat het moeilijk zou zijn kinderen op een informele manier het nieuws te laten lezen of bekijken als

zij alleen zouden zijn met een vreemde proefleider.

De kinderen werden in groepjes van zes of acht uit hun klaslokaal gehaald. De kinderen die aan de televisieconditie waren toegewezen werden door een proefleider naar de 'televisieruimte' gebracht; de overige leerlingen gingen met een andere proefleider mee naar de 'leesruimte'. In beide locaties zaten kinderen aan tafels die ver uit elkaar gezet waren. In de televisieruimte zaten de kinderen ongeveer twee meter verwijderd van een 40cm-kleurentelevisie, die op ooghoogte was geplaatst.

Om mogelijke school- of groepseffecten te vermijden, waren alle vier de experimentele condities vertegenwoordigd in iedere school en klas. Per klas namen maximaal 16 kinderen aan het experiment deel. Om te voorkomen dat kinderen elkaar op de hoogte stelden van wat er in het experiment van hen werd verwacht, werden de experimentele behandelingen van kinderen uit dezelfde klas in één ochtend of middag afgerond. De kinderen uit de klas die toegewezen waren aan de televisie- of tekstconditie zonder toetsverwachting namen deel aan de eerste sessie. De overige kinderen, die waren toegewezen aan een televisie- of tekstconditie met toetsverwachting, werden getest in een aansluitende tweede sessie. Een sessie duurde ongeveer 50 minuten.

In de condities zonder toetsverwachting werd kinderen verteld: "We gaan vandaag eens iets anders, iets bijzonders doen. Ik moet alleen nog wat voor jullie voorbereiden, en dat kan ik nu pas doen, omdat ik nu pas weet hoe jullie heten." In de televisieconditie vervolgde de proefleider: "In de tussentijd kunnen jullie wat naar de televisie kijken. Ik heb hier een videoband met een aflevering van het Jeugdjournaal. Daar kunnen jullie naar kijken terwijl ik bezig ben." In de tekstconditie zei de proefleider: "In de tussentijd kunnen jullie wat lezen. Ik heb hier een paar stukjes uit een kinderkrant. Die kunnen jullie lezen terwijl ik bezig ben." Op een onnadrukkelijke manier werd aan de kinderen gevraagd om niet met elkaar te praten, omdat de proefleider zich anders moeilijk zou kunnen concentreren.

In de condities met toetsverwachting werd kinderen gezegd: "We gaan vandaag een testje doen." In de televisieconditie zei de proefleider: "Ik heb hier een videoband met een afleve-

ring van het Jeugdjournaal, die we gaan bekijken." In de tekstconditie werd gezegd: "Ik heb hier een paar stukjes uit een kinderkrant, die jullie moeten lezen." Vervolgens werd de instructie gegeven: "Jullie moeten goed opletten bij wat jullie straks gaan zien (of: gaan lezen), want na afloop krijgen jullie er vragen over."

Terwijl de expositietijd voor de kinderen die het nieuws via de televisie zagen steeds 11 minuten bedroeg, werd de kinderen in de tekstconditie toegestaan het nieuws in hun eigen tempo te lezen. Er werd van afgezien om de expositietijd voor de televisie- en tekstcondities gelijk te stellen, omdat een vooronderzoek had uitgewezen dat de meeste kinderen meer dan 11 minuten nodig hadden om de vijf nieuwsberichten te lezen. Gelijktelling van de expositietijd zou daarom de tekstconditie onrecht aandoen en een zinvolle mediumvergelijking in de weg staan. De gemiddelde leestijd bedroeg 12.59 minuten.

Onmiddellijk na vertoning van de televisieversie kregen de kinderen in de televisieconditie de vragenlijst voor de eerste afhankelijke variabele gepresenteerd. In de tekstconditie waren de kinderen niet altijd op hetzelfde moment klaar met lezen. Hen werd gevraagd een hand op te steken als zij klaar waren en vervolgens werd hen de eerste vragenlijst overhandigd. De eerste maat die werd afgenomen betrof de mentale inspanning die kinderen in het bekijken of lezen van de nieuwsonderwerpen hadden geïnvesteerd. Vervolgens werd aan de hand van een 'cued-recall test' vastgesteld wat kinderen van de onderwerpen hadden onthouden. Beide maten werden voorafgegaan door een schriftelijke uitleg. Het leesvaardigheidsniveau van de kinderen was reeds drie weken voorafgaand aan het experiment vastgesteld.

2.5 Meetinstrumenten

Begrijpend lezen

Het leesvaardigheidsniveau van de kinderen werd gemeten met behulp van gestandaardiseerde tests voor begrijpend lezen, ontwikkeld door het CITO. Verschillende tests werden gebruikt voor kinderen uit groep 6 (CITO, 1981) en groep 8 (CITO, 1991). Beide toetsen bestonden uit vijf teksten waarover in totaal 25 meerkeuzevragen werden gesteld. In de vragen bij elke tekst kregen kinderen de opdracht het centrale thema aan te wijzen, verschillende gedeel-

ten van de tekst met elkaar te verbinden en gevolgtrekkingen te maken uit de gegeven informatie. Cronbachs alfa (K-R 20) was .85 voor groep 6 en .80 voor groep 8.

Mentale inspanning

De hoeveelheid mentale inspanning die kinderen in het bekijken en lezen van het nieuws investeerden, werd gemeten met behulp van een uitgebreide versie van Salomons (1984) schaal voor mentale inspanning, ontwikkeld door Beentjes, Vooijs en Van der Voort (1993). De uitgebreide vragenlijst bestond uit 18 uitspraken waarin de mate van inspanning tijdens het verwerken van het nieuws werd uitgedrukt. Bijvoorbeeld: "Toen ik aan het televisiekijken was (of: toen ik aan het lezen was), wilde ik alles begrijpen." Kinderen reageerden op iedere uitspraak aan de hand van een vierpunts-schaal, die liep van "helemaal mee eens" tot "helemaal niet mee eens". Cronbachs alfa was .79 in de televisieconditie en .81 in de tekstconditie.

Herinnering

De retentie van de vijf berichten werd gemeten aan de hand van een schriftelijke kennistoets bestaande uit 56 open vragen. Er werd gebruik gemaakt van zogenaamde 'cued-recall'-vragen, in plaats van een vrije herinneringsmaat waarbij kinderen spontaan opschrijven wat zij nog weten, omdat vrije herinneringstoetsen de feitelijke kennis van proefpersonen kunnen onderschatten (Berry, 1983; Woodall, Davis & Sahin, 1983). Van het gebruik van een meerkeuzetoets werd afgezien, omdat we de daadwerkelijke herinnering van het nieuws wilden meten en niet herkenning.

Voor ieder nieuwsbericht werden open vragen opgesteld die betrekking hadden op elk van de vijf voornaamste elementen waaruit de meeste nieuwsonderwerpen bestaan: gebeurtenis, plaats, belangrijkste personen, oorzaak en gevolg (Findahl & Høijer, 1985). Daarnaast werden voor ieder bericht vragen over details opgenomen. In de kennistoets werden alleen vragen gesteld over informatie die was opgenomen in de tekst van de gedrukte nieuwsberichten (en het gesproken commentaar). Er werden dus geen vragen gesteld over visuele informatie die niet verbaal was weergegeven. De formulering van een aantal vragen werd bijgesteld

naar aanleiding van een vooronderzoek uitgevoerd bij kinderen uit de groepen 6 en 8 van vier scholen die niet aan het experiment deelnamen ($N = 148$). Cronbachs alfa (K-R 20) voor de definitieve toets was .93.

In de televisieberichten werd een gedeelte van de verbaal weergegeven informatie aangevuld met beelden die min of meer dezelfde informatie bevatten. Om een analyse te kunnen maken van de bijdrage die deze redundante visuele informatie aan de herinnering levert, werden de vragen uit de kennistoets ingedeeld in twee typen: (a) vragen over informatie die alleen verbaal was weergegeven, in de tekst- en in de televisieversie, en (b) vragen over informatie die, geheel of gedeeltelijk, ook in beelden was weergegeven in de televisieversie. De 56 vragen werden door twee onafhankelijke beoordelaars ingedeeld in de categorieën 'puur verbaal' en 'gevisualiseerd'. De tussenbeoordelaars-overeenstemming, gemeten met behulp van Cohens kappa, was .96. Vijfentwintig vragen werden gecategoriseerd als 'puur verbaal' (Cronbachs alfa = .89) en 21 vragen als 'gevisualiseerd' (Cronbachs alfa = .84).

3 Resultaten

Data-controle liet zien dat de verdelingen voor de mentale inspannings- en herinneringsscores voldeden aan de aannames voor variantieanalyse. Voor geen enkele maat werden hoofdeffecten van sekse of interactie-effecten met sekse gevonden. Sekse werd daarom als variabele buiten de analyses gehouden. In de variantieanalyses werden de proefpersonen die samen een blok van vier vormden (en die gematched waren op leesvaardigheid) behandeld alsof zij één proefpersoon waren die meerdere herhaalde metingen heeft ondergaan (Kirk, 1968; Tabachnick & Fidell, 1989). De factoren 'medium' en 'toetsverwachting' werden dus in de analyses behandeld als binnenproefpersonen-, ofwel 'binnen-blokken'-factoren. Alle effecten werden getoetst onder gebruikmaking van een alfaniveau van .05.

3.1 Mentale inspanning

Een 2 (tekst vs. televisie) x 2 (met vs. zonder toetsverwachting) x 2 (groep 6 vs. groep 8) x 2 (matige vs. goede lezers)-variantie-analyse

Tabel 1*Gemiddelde mentale inspanning van kijkers en lezers als een functie van toetsverwachting en leeftijdsgroep*

groep	Toetsverwachting							
	Ja				Nee			
	TV		Tekst		TV		Tekst	
	M	SD	M	SD	M	SD	M	SD
6	2.94	0.50	2.85	0.34	2.61	0.43	2.93	0.38
8	2.97	0.32	3.07	0.27	2.60	0.36	2.69	0.37
M	2.96	0.41	2.96	0.32	2.61	0.39	2.81	0.52

Noot Minimum score is 1 en maximum score is 4.

werd uitgevoerd op de scores voor mentale inspanning, met medium en toetsverwachting als binnen-proefpersonen-factoren en leeftijdsgroep en niveau voor begrijpend lezen als tussen-proefpersonen-factoren. Tabel 1 vermeldt het gemiddelde van de mentale inspanningsscores (totaalscores per proefpersoon gedeeld door het aantal items) voor de televisie- en tekstconditie, per leeftijdsgroep en voor de condities met en zonder toetsverwachting. Het hoofdeffect voor toetsverwachting was significant, $F(1, 34) = 18.98, p < .001$; de proportie verklaarde variantie was $\eta^2 = .36$, waarbij $\eta^2 = SS_{\text{effect}} / SS_{\text{effect}} + SS_{\text{error}}$ (Cohen, 1973, p. 108; 1988; Tabachnick & Fidell, 1989, p. 444). Conform de bedoeling, rapporteerden de kinderen die te voren wisten dat zij getoetst zouden worden dat zij zich meer hadden ingespannen dan de kinderen die vooraf niet wisten dat er een test zou volgen. Er waren geen significante hoofdeffecten voor medium, leeftijdsgroep of niveau voor begrijpend lezen. Overeenkomstig de voorspelling, leidde de aankondiging van een toets bij de kijkers tot een grotere verhoging van de mentale inspanning dan bij de lezers. Dit interactie-effect tussen medium en

toetsverwachting was echter niet significant, $F(1, 34) = 2.39, p < .13$. Er was echter wel een significante interactie tussen toetsverwachting en leeftijdsgroep: de mentale inspanning van kinderen uit groep 8 werd door de aankondiging van een toets sterker verhoogd dan bij kinderen uit groep 6, $F(1, 34) = 4.93, p < .03, \eta^2 = .13$. Er waren geen andere significante interactie-effecten.

3.2 Herinnering

Een 2 (tekst vs. televisie) x 2 (met vs. zonder toetsverwachting) x 2 (groep 6 vs. groep 8) x 2 (matige vs. goede lezers)-variantie-analyse werd uitgevoerd op de totaalscores op de herinneringstoets, met medium en toetsverwachting als binnen-proefpersonen-factoren en leeftijdsgroep en niveau voor begrijpend lezen als tussen-proefpersonen-factoren. Tabel 2 vermeldt de gemiddelde proporties goede antwoorden van de kijkers en lezers, per leeftijdsgroep en voor de condities met en zonder toetsverwachting. Zoals voorspeld, onthielden de kinderen die het nieuws via de televisie hadden gezien ($M = .51, SD = .19$) significant meer dan de kinderen die de tekstversies hadden ge-

Tabel 2*Gemiddelde herinneringsscores van kijkers en lezers als een functie van toetsverwachting en leeftijdsgroep*

groep	Toetsverwachting							
	Ja				Nee			
	TV		Tekst		TV		Tekst	
	M	SD	M	SD	M	SD	M	SD
6	.44	.19	.32	.19	.38	.17	.31	.16
8	.64	.12	.51	.20	.56	.15	.53	.22
M	.55	.19	.42	.21	.47	.19	.43	.22

Noot De scores zijn gemiddelde proporties goede antwoorden.

lezen ($M = .42$, $SD = .21$), $F(1, 34) = 10.62$, $p < .003$, $\eta^2 = .24$.

Kinderen uit groep 8 ($M = .56$, $SD = .18$) onthielden significant meer van het nieuws dan kinderen uit groep 6 ($M = .36$, $SD = .18$), $F(1, 34) = 61.07$, $p < .001$, $\eta^2 = .64$. Kinderen die te voren wisten dat zij getoetst zouden worden, presteerden iets beter op de kennistoets dan de kinderen zonder toetsverwachting, maar dit effect was niet significant, $F(1, 34) = 2.87$, $p < .10$, en het was niet van toepassing op de kinderen uit de tekstcondities. Conform de voorspelling, was de kennis van de kijkers meer bij een toetsaankondiging gebaat dan de kennis van de lezers, maar ook deze trend was niet significant, $F(1, 34) = 2.07$, $p < .16$. Anders dan verwacht, verschilde de relatieve effectiviteit van televisie- en gedrukt nieuws niet per leeftijdsgroep, $F(1, 34) = 0.04$, $p < .85$. Zowel kinderen uit groep 6 als kinderen uit groep 8 onthielden meer van het televisienieuws dan van het gedrukte nieuws en dit leervoordeel van de televisie was bij beide leeftijdsgroepen ongeveer even groot. Er waren geen andere significante interactie-effecten.

Tabel 3 bevat de gemiddelde proporties goede antwoorden op de kennistoets voor de televisie- en tekstconditie, uitgesplitst naar niveau van begrijpend lezen. De goede lezers onthielden meer van het nieuws dan de matige lezers, $F(1, 34) = 50.11$, $p < .001$, $\eta^2 = .59$. Het leervoordeel van televisie- boven gedrukt

nieuws was groter bij de matige lezers dan bij de goede lezers, maar tegen onze verwachting in was dit interactie-effect tussen medium en niveau van begrijpend lezen niet statistisch significant, $F(1, 34) = 3.35$, $p < .08$.

Om te controleren of het voordeel van televisie- boven gedrukt nieuws opging voor elk van de vijf nieuwsberichten die in het onderzoek werden gebruikt, werd een variantie-analyse uitgevoerd met medium en toetsverwachting als 'binnen-blokken'-factoren en de herinneringsscores per nieuwsbericht als additionele binnen-proefpersonen-factor. Tabel 4 vermeldt per nieuwsverhaal de gemiddelde proporties goede antwoorden voor de televisie- en tekstconditie. Er was geen interactie-effect tussen de vijf nieuwsonderwerpen en de factor medium, $F(4, 148) = 1.19$, $p < .32$, hetgeen betekent dat voor elk van de vijf berichten gold dat de televisieversie beter werd onthouden dan de tekstversie.

Om te onderzoeken wat de bijdrage aan de herinnering was van de aan de verbale boodschappen gerelateerde visuele informatie uit de televisieberichten, werd een variantie-analyse uitgevoerd met medium en toetsverwachting als 'binnen-blokken'-factoren en als additionele binnen-proefpersonen-factor de herinneringsscores voor de 'puur verbale' en 'gevisualiseerde' informatie. Blijkens een significant interactie-effect tussen medium en deze twee categorieën informatie (zie Figuur 1), deed het leervoordeel van televisie boven

Tabel 3

Gemiddelde herinneringsscores van kijkers en lezers als een functie van niveau van begrijpend lezen

	TV		Tekst		Gemiddeld	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Begrijpend Lezen						
Goed	.58	.13	.54	.16	.56	.13
Matig	.44	.15	.31	.16	.38	.13

Noot. De scores zijn gemiddelde proporties goede antwoorden.

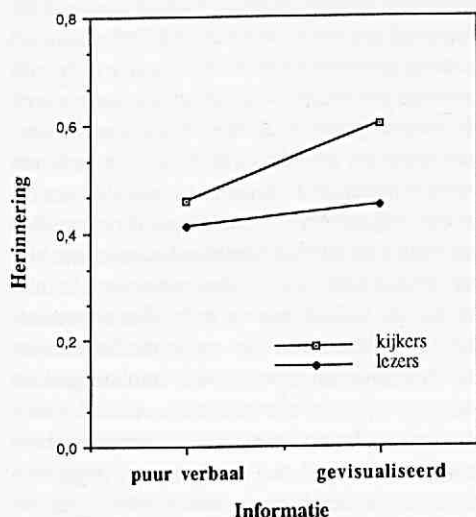
Tabel 4

Gemiddelde herinneringsscores van kijkers en lezers per nieuwsonderwerp

Onderwerp	1		2		3		4		5	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
TV	.52	.20	.54	.27	.51	.23	.53	.25	.47	.21
Tekst	.42	.24	.48	.25	.39	.23	.42	.27	.40	.25
<i>M</i>	.47	.23	.51	.26	.45	.23	.48	.27	.44	.23

Noot. De scores zijn gemiddelde proporties goede antwoorden.

tekst zich het sterkst voor bij de informatie die in de televisieberichten was gevisualiseerd, $F(1, 37) = 27.21, p < .001, \eta^2 = .42$. Op de kennisvragen over informatie die alleen verbaal werd overgedragen, scoorden kijkers ($M = .46, SD = .20$) slechts iets hoger dan lezers ($M = .41, SD = .22$), terwijl er een veel groter verschil bestond tussen kijkers ($M = .60, SD = .20$) en lezers ($M = .45, SD = .22$) op de kennisvragen over informatie die op de televisie was gevisualiseerd.



Figuur 1. Gemiddelde herinneringsscores van kijkers en lezers voor puur verbale en gevisualiseerde informatie

4 Discussie

Terwijl eerdere experimentele vergelijkende studies lieten zien dat volwassenen gedrukt nieuws beter onthouden dan televisienieuws, werd in de huidige studie bij kinderen een tegenovergesteld resultaat gevonden. In overeenstemming met de hoofdhypothese (H_1), onthielden kinderen nieuwsberichten van de televisie beter dan vergelijkbare tekstversies. Het gedrukte nieuws werd het slechtst onthouden, ondanks het feit dat de tekstconditie op twee manieren werd bevoordeeld. Ten eerste werden de lezers gemiddeld twee minuten langer aan het stimulusmateriaal blootgesteld dan de kijkers. Daarnaast werden de allerzwakste lezers buiten het experiment gehouden, omdat verwacht werd dat de opname van deze groep de resultaten van de tekstconditie zou drukken.

De superioriteit van televisie als medium om nieuwsinformatie aan kinderen over te dragen had een algemener karakter dan verwacht. De bevinding dat kinderen televisienieuws beter onthouden dan gedrukt nieuws bleef niet beperkt tot bepaalde subgroepen van kinderen, maar gold voor het volledige leeftijdsbereik dat werd onderzocht, ongeacht het leesvaardigheidsniveau van de kinderen. Daarnaast bleef de bevinding dat televisie voor kinderen het 'beste' nieuwsmedium is, niet beperkt tot de onnatuurlijke situatie waarin nieuwsverhalen worden geconsumeerd in de wetenschap dat men later over de informatie zal worden ondervraagd, maar gold zij ook voor de meer natuurlijke situatie waarin het nieuws zonder toetsverwachting wordt geconsumeerd. De resultaten van dit onderzoek zijn 'goed nieuws' voor kinderen, omdat zij het medium dat als beste naar voren kwam aanzienlijk frequenter gebruiken dan de krant.

Hoewel zoals voorzien in H_2 , het leervoordeel van televisie- boven gedrukt nieuws bij de zwakke lezers groter was dan bij de goede lezers, was dit verschil niet significant. Een mogelijke reden waarom geen significante interactie tussen leesvaardigheid en medium werd gevonden, is dat de allerzwakste lezers buiten het experiment werden gehouden. Een tweede mogelijke verklaring kan worden gezocht in het geringe moeilijkheidsniveau van de gebruikte nieuwsverhalen. De teksten waren helder geschreven en sloten goed aan bij het begripsniveau van kinderen van 10 tot 12 jaar. Het is denkbaar dat het leesvaardigheidsniveau van kinderen pas interacteert met de relatieve effectiviteit van televisie- en gedrukt nieuws indien de teksten minder eenvoudig te begrijpen zijn. Zwakke lezers ondervinden bij het verwerken van informatie met een hoog moeilijkheidsniveau extra problemen, een handicap die in de televisieconditie deels overwonnen kan worden door de steun die begeleidende beelden bij de verwerking van de verbale informatie kunnen bieden.

Het geringe moeilijkheidsniveau van de gebruikte teksten kan mogelijk tevens verklaren waarom, anders dan voorspeld (H_3), het leervoordeel van televisie- boven gedrukt nieuws bij jonge kinderen (groep 6) niet hoger uitviel dan bij oudere kinderen (groep 8). Weliswaar hadden de jongere kinderen vergeleken bij de

oudere kinderen meer moeite om zowel het gedrukte als het televisienieuws te onthouden. Maar omdat de teksten ook voor de jongste kinderen goed te begrijpen waren, werden zij door hun geringere leesvaardigheid niet zodanig gehinderd dat het leernadeel van tekst ten opzichte van televisie bij hen sterker uitviel dan bij de oudere kinderen.

In vergelijking met de conditie zonder toetsverwachting, leidde de aankondiging van een toets conform de bedoeling tot een significante verhoging van de mentale inspanning van de kijkers en de lezers. In overeenstemming met Salomon en Leigh's (1984) hypothese (H_4), werd vooral de mentale inspanning van de kijkers door de aankondiging van een natoets verhoogd, maar ook dit interactie-effect was niet significant. Dat de door een toetsaankondiging opgeroepen verhoging van de mentale inspanning bij kijkers niet veel hoger was dan bij lezers, kan het gevolg zijn van het feit dat de kijkers ook zonder toetsverwachting reeds een relatief hoog niveau van mentale inspanning vertoonden (zij hadden een gemiddelde score van 2.61 op een schaal van 1 tot 4). Ondanks onze pogingen om de conditie zonder toetsverwachting zo min mogelijk te laten afwijken van de thuissituatie, was de mentale inspanning in deze conditie mogelijk toch hoger dan in de thuissituatie, wellicht omdat de kinderen uit de klas werden gehaald om iets onbekends te gaan doen en in kleine groepjes werden getest. In het onderzoek van Beentjes et al. (1993) werden halve schoolklassen tegelijkertijd getest; Salomon en Leigh verschaffen hieromtrent geen informatie.

Onze voorspelling dat het leervoordeel van televisie boven gedrukt nieuws het hoogst zou uitvallen bij kinderen die wisten dat zij getoetst zouden worden (H_5), was gebaseerd op de aanname dat een toetsverwachting vooral de mentale inspanning van de kijkers verhoogt (H_4). Omdat deze aanname niet bleek op te gaan, is het begrijpelijk dat het leervoordeel van televisie boven tekst in de conditie met toetsverwachting niet significant hoger was dan in de conditie zonder toetsverwachting.

De voorspelling dat kinderen televisienieuws beter onthouden dan gedrukt nieuws was gedeeltelijk gebaseerd op de aanname dat kinderen gebaat zijn bij de extra geheugensteun die de televisie via haar beelden biedt, mits de

beeldinformatie gerelateerd is aan het gesproken commentaar. Deze veronderstelling kreeg in ons onderzoek steun van de bevinding dat het leervoordeel van televisie boven tekst het meest geprononceerd was bij die elementen uit de verbale informatie die op de televisie door inhoudelijk aansluitende beelden werden ondersteund.

De gevonden superioriteit van televisie boven gedrukt nieuws is dus niet alleen het gevolg van verschillen die inherent zijn aan beide communicatiemediën en de manier waarop kinderen informatie uit deze media verwerken. Zij is mede te danken aan de wijze waarop de makers van het Jeugdjournaal het medium televisie hebben gebruikt. In de gebruikte onderwerpen werd het grote voordeel van televisie om verbale informatie visueel te ondersteunen effectief benut, terwijl tegelijkertijd de nadelen die vaak aan televisienieuwsuitzendingen kleven zoveel mogelijk werden vermeden. In feite hebben de makers van de gebruikte nieuwsberichten voldaan aan een reeks aanbevelingen die Robinson en Levy (1986) hebben gedaan om een effectieve overdracht van televisienieuws te bevorderen. De nieuwsberichten werden in een relatief rustig tempo gepresenteerd, de onderwerpen werden helder van elkaar gescheiden, waar mogelijk werd een 'human-interest'- invalshoek bij de verhalen gebruikt, de nieuwsinformatie werd helder verwoord, technische en gespecialiseerde termen werden uitgelegd en waar gewenst werd gebruik gemaakt van grafische ondersteuning.

Onze indruk is dat het Jeugdjournaal, door de doelmatige manier waarop het medium televisie wordt gebruikt, herinneringstechnisch beter in elkaar steekt dan het 'volwassen' journaal dat de zojuist besproken aanbevelingen van Robinson en Levy (1986) minder stringent ter harte lijkt te nemen. Het is daarom nog maar de vraag of de bij volwassenen aangetroffen superioriteit van gedrukt nieuws ook van toepassing is wanneer hun herinnering van berichten uit het Jeugdjournaal met gedrukte versies wordt vergeleken. Het is geenszins uitgesloten dat volwassenen, ondanks hun hogere leesvaardigheid, van nieuwsverhalen uit het Jeugdjournaal evenveel onthouden als van vergelijkbare gedrukte verhalen, en wellicht wordt het Jeugdjournaal zelfs beter onthouden. Evenzeer is het de vraag of de bij kinderen aangetroffen

superioriteit van televisie ook geldt wanneer hun herinnering van het 'volwassen' televisie-journaal met gedrukt nieuws wordt vergeleken. In een vervolgonderzoek, dat thans gaande is, worden bovengenoemde vragen onderzocht door na te gaan hoeveel volwassenen en kinderen onthouden van televisie- en tekstversies van nieuwsverhalen die door het Jeugdjournaal en het 'volwassen' journaal over dezelfde onderwerpen zijn vervaardigd.

Terwijl de voordelen die televisie bij de informatieoverdracht biedt in de gebruikte nieuwsverhalen ten volle zijn uitgebuit, kan men zich afvragen of de voordelen die de krant biedt in het onderzoek niet zijn onderbenut. De gedrukte nieuwsversies werden, evenals in de eerdere vergelijkende studies bij volwassenen, direct afgeleid van de televisieteksten. Omwille van de experimentele vergelijkbaarheid waren de gedrukte teksten identiek aan het gesproken televisiecommentaar. Hoewel de gedrukte verhalen ook zonder kennis van de begeleidende beelden als op-zichzelf-staande tekst uitstekend te begrijpen waren, is het in beginsel denkbaar dat het medium tekst in onze studie enigszins tekort werd gedaan, omdat de verhalen niet speciaal voor de krant waren geschreven. Dit probleem deed zich ook voor in de mediumvergelijkende experimenten die bij volwassenen werden uitgevoerd. Maar in dat onderzoek stond de mogelijke onderbenutting van het medium tekst de interpretatie van de gevonden mediumverschillen niet in de weg, omdat het gedrukte nieuws het beste werd onthouden. Omdat kinderen zich echter het meest van het televisienieuws herinnerden, dient nader te worden onderzocht of het gevonden mediumverschil niet mede is toe te schrijven aan een onderbenutting van het medium tekst. In een vervolgonderzoek zullen wij daarom nagaan of televisie voor kinderen ook het meest effectieve nieuwsmedium is wanneer nieuwsberichten van de televisie behalve met letterlijk uitgeschreven commentaren vergeleken worden met artikelen die door verschillende jeugd-journalisten over dezelfde onderwerpen zijn geschreven. Wij verwachten echter dat ook uit een dergelijke vergelijking de televisie als het meest effectieve nieuwsmedium voor kinderen uit de bus zal komen.

Literatuur

- Anderson, B., Mead, N., & Sullivan, S. (1986). *Television: What do National Assessment results tell us?* Princeton, NJ: Educational Testing Service.
- Beagles-Ross, J., & Gat, I. (1983). Specific impact of radio and television on children's story comprehension. *Journal of Educational Psychology*, 75, 128-137.
- Beentjes, J. W. J., & Voort, T. H. A. van der (1991a). Recall and language use in retellings of televised and printed stories. *Poetics*, 20, 91-104.
- Beentjes, J. W. J., & Voort, T. H. A. van der (1991b). Children's written accounts of televised and printed stories. *Educational Technology Research and Development*, 39(3), 15-26.
- Beentjes, J. W. J., & Voort, T. H. A. van der (1993). Television viewing versus reading: Mental effort, retention, and inferential learning. *Communication Education*, 42, 191-205.
- Beentjes, J. W. J., Vooijs, M. W., & Voort, T. H. A. van der (1993). Children's recall of televised and printed news as a function of test expectation. *Journal of Educational Television*, 19, 5-13.
- Berry, C. (1983). Learning from television news: A critique of the research. *Journal of Broadcasting*, 27, 359-370.
- CITO (1981). *Begrijpend lezen leerjaar 3, 4 en 5 basisonderwijs*. Arnhem: CITO.
- CITO (1991). *Eindtoets basisonderwijs*. Arnhem: CITO.
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor anova designs. *Educational and Psychological Measurement*, 33, 107-112.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Comstock, G., & Paik, H. (1991). *Television and the American child*. New York: Academic Press.
- Craik, F. I. M. (1979). Human memory. *Annual Review of Psychology*, 30, 63-102.
- Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research Vol II* (pp. 512-538). New York: Longman.
- DeFleur, M. L., Davenport, L., Cronin, M., & DeFleur, M. (1992). Audience recall of news stories presented by newspaper, computer, television, and radio. *Journalism Quarterly*, 69, 1010-1022.
- Drew, D. G., & Grimes, T. (1987). Audio-visual redundancy and TV news recall. *Communication Research*, 14, 452-461.
- Drew, D. G., & Reeves, B. B. (1980). Children and television news. *Journalism Quarterly*, 57, 45-54.

- Facorro, L. B., & DeFleur, M. L. (1993). A cross-cultural experiment on how well audiences remember news stories from newspaper, computer, television, and radio. *Journalism Quarterly*, 70, 585-601.
- Findahl, O., & Høijer, B. (1985). Some characteristics of news memory and comprehension. *Journal of Broadcasting and Electronic Media*, 29, 379-396.
- Furnham, A., & Gunter, B. (1985). Sex, presentation mode and memory for violent and non-violent news. *Journal of Educational Television*, 11, 99-105.
- Furnham, A., Gunter, B., & Green, A. (1990). Remembering science: The recall of factual information as a function of the presentation mode. *Applied Cognitive Psychology*, 4, 203-212.
- Gibbons, J., Anderson, D. R., Smith, R., Field, D. E., & Fischer, C. (1986). Young children's recall and reconstruction of audio and audiovisual narratives. *Child Development*, 57, 1014-1023.
- Graber, D. A. (1990). Seeing is remembering: How visuals contribute to learning from television news. *Journal of Communication*, 40(3), 134-155.
- Greenfield, P. M., & Beagles-Roos, J. (1988). Television vs. radio: The cognitive impact on different socioeconomic and ethnic groups. *Journal of Communication*, 38(2), 71-92.
- Grimes, T. (1991). Mild auditory-visual dissonance in television news may exceed viewer attentional capacity. *Human Communication Research*, 18, 268-298.
- Gunter, B. (1987). *Poor reception: Misunderstanding and forgetting broadcast news*. Hillsdale, NJ: Erlbaum.
- Gunter, B. (1991). Responding to news and public affairs. In J. Bryant & D. Zillmann (Eds.), *Responding to the screen: Reception and reaction processes* (pp. 229-260). Hillsdale, NJ: Erlbaum.
- Gunter, B., & Furnham, A. (1986). Sex and personality differences in recall of violent and non-violent news from three presentation modalities. *Personality and Individual Differences*, 7, 829-837.
- Gunter, B., Furnham, A., & Gietson, G. (1984). Memory for the news as a function of the channel of communication. *Human Learning*, 3, 265-271.
- Gunter, B., Furnham, A., & Leese, J. (1986). Memory for information from a party political broadcast as a function of the channel of communication. *Social Behaviour*, 1, 135-142.
- Hayes, D. S., Kelly, S. B., & Mandel, M. (1986). Media differences in children's story synopses: Radio and television contrasted. *Journal of Educational Psychology*, 78, 341-346.
- Kirk, R. E. (1968). *Experimental design: Procedures for the behavioral sciences*. Belmont, CA: Brooks/Cole Publishing Company.
- Kozma, R. B. (1991). Learning with media. *Review of Educational Research*, 61, 179-211.
- Lil, J. E. van (1989). Media use by children and young people. *EBU Review*, 40(4), 23-27.
- Meringoff, L. K. (1980). Influence of the medium on children's story apprehension. *Journal of Educational Psychology*, 72, 240-249.
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, 61, 179-211.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt.
- Pezdek, K., Lehrer, A., & Simon, S. (1984). The relationship between reading and cognitive processing of television and radio. *Child Development*, 54, 1015-1023.
- Reese, S. D. (1984). Visual-verbal redundancy effects on television news learning. *Journal of Broadcasting*, 28, 79-87.
- Robinson, J. P., & Davis, D. K. (1990). Television news and the informed public: An information-processing approach. *Journal of Communication*, 40(3), 106-119.
- Robinson, J. P., & Levy, M. R. (1986). *The main source: Learning from television news*. Beverly Hills, CA: Sage.
- Roper Reports (1995). *America's watching: Public attitudes toward television*. New York: Roper Starch Worldwide Inc.
- Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, 74, 647-658.
- Salomon, G., & Leigh, T. (1984). Predispositions about learning from print and television. *Journal of Communication*, 34(2), 119-135.
- Stauffer, J., Frost, R., & Rybolt, W. (1981). Recall and learning from broadcast news: Is print better? *Journal of Broadcasting*, 25, 253-262.
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper & Row.
- Wember, B. (1976). *Wie informiert das Fernsehen?* München, Duitsland: List.
- Wicks, R. H., & Drew, D. G. (1991). Learning from news: Effects of message consistency and medium on recall and inference making. *Journalism Quarterly*, 68, 155-164.

- Wilson, C. E. (1974). The effect of medium on loss of information. *Journalism Quarterly*, 51, 111-115.
- Woodall, W. G., Davis, D. K., & Sahin, H. (1983). From the boob tube to the black box: Television news comprehension from an information processing perspective. *Journal of Broadcasting*, 27, 1-23.

Manuscript aanvaard 11-1-1996

Auteurs

J. H. Walma van der Molen is ontwikkelingspsychologe en als Onderzoeker in Opleiding verbonden aan de Vakgroep Algemene Pedagogiek, Rijksuniversiteit Leiden. Zij doet promotie-onderzoek naar het onthouden van televisie- en gedrukt nieuws door kinderen.

T. H. A. van der Voort is als hoogleraar empirische pedagogiek verbonden aan de Rijksuniversiteit Leiden en is wetenschappelijk directeur van ISED.

Correspondentie-adres: J. H. Walma van der Molen, Sectie Kind & Media, Rijksuniversiteit Leiden, Postbus 9555, 2300 RB Leiden.

Abstract

Learning effects of televised and printed news for children: A media comparison study.

J. H. Walma van der Molen & T. H. A. van der Voort.
Pedagogische Studiën, 1996, 73, 459-473.

An experiment was conducted (a) to compare children's recall of news information presented either audio-visually or in print, and (b) to establish whether the relative effectiveness of both media in conveying the news is dependent on children's level of reading proficiency and test expectation. A sample of 152 fourth and sixth graders was presented with a sequence of five children's news stories, either in their original televised form or in a printed version. In each condition, half of the children were led to expect a retention test, whereas the other children were not. The results of a cued-recall test indicated that children who had watched the news on television remembered more than those who had received the same news in print, regardless of their reading proficiency or expectation of a retention test.

Boekbespreking

H. A. A. van Eerde

Kwantiwijzer. Diagnostiek in reken-wiskundeonderwijs

Academisch proefschrift Universiteit Utrecht. Uitgeverij
Zwijsen, Tilburg 1996, 437 pagina's, ISBN
90 276 3673 7

M. van den Heuvel-Panhuizen

Assessment and realistic mathematics education

Freudenthal Instituut, Utrecht 1996, 305 pagina's, ISBN
90 393 1333 4

A. M. van de Rijt

Vorbereidende rekenvaardigheid bij kleuters. De ontwikkeling van rekenvaardigheids-schalen en een onderzoek naar de invloed van een programma

Academisch proefschrift Universiteit Utrecht. Graviant
Educatieve Uitgaven, Doetinchem 1996, 165 pagina's,
ISBN 90 751 2908 4

In het aprilnummer van het tijdschrift 'Het Jonge Kind' (1996) vraagt een groep pedagogen en psychologen – en sommigen van hen zijn nauw bij de opvoeding van kleuters betrokken – in een ingezonden mededeling de kritische aandacht van de lezer voor 'een onwijs idee'. Bedoeld wordt het plan van de huidige staatssecretaris om een zogenoemde entreetoets voor vierjarigen te ontwikkelen. Met die toets zouden volgens de bewindsvrouw de 'zorgbehoeften' van het jonge kind in kaart gebracht kunnen worden zodat achterstanden effectief kunnen worden bestreden. De argumenten die tegen dit plan worden ingebracht zijn deels toetstechnisch en grotendeels fundamenteel van aard zodat de beleidsmakers waarschijnlijk alle zeilen zullen moeten bijzetten om het bekritiseerde initiatief overtuigend te kunnen verdedigen. De discussie over de betekenis, de voor- en de nadelen van toetsen voor jonge kinderen is pedagogisch en didactisch van belang. Dat moge ook blijken uit het feit dat er in het voorjaar liefst drie dissertaties werden verdedigd (alle drie in Utrecht), waarin diepgaand aandacht wordt besteed aan de diagnostiek van de ontwikkeling van cognitieve – in het bijzonder wiskundige vaardigheden – van jonge kinderen. De studies hebben voor die discussie waardevolle gegevens en informatie

opgeleverd. De drie onderzoeksters laten namelijk, ieder op eigen wijze, zien hoe een sprekend en betrouwbaar beeld kan worden verkregen van de ontwikkeling van cognitieve vaardigheden van jonge kinderen (4 – 7 jaar). In het bijzonder hebben de onderzoeksters studie gemaakt van de ontwikkeling van inzichten en vaardigheden die als de basis worden beschouwd om met succes het reken-wiskundeonderwijs te kunnen volgen. Het gaat hier niet om entreetoetsen, maar om interactieve toetsen die aan kinderen ruimte laten (zeker wat twee van de drie studies betreft) om tot eigen vondsten te komen en die een basis bieden voor diagnostisch, preventief onderwijs.

De studie van Van Eerde kan getypeerd worden als een kritische en legitimerende terugblik op een project waarin systematisch werd gewerkt aan de ontwikkeling van een diagnostisch instrument op basis van geëxpliciteerde theoretische inzichten. De werkwijzen en inzichten worden post-hoc aan een kritische analyse onderworpen.

Van den Heuvel-Panhuizen wil op basis van theoretische beschouwingen en analyses laten zien hoe het begrip 'assessment' uitgaande van de realistische wiskundige vakdidactiek uitgewerkt moet worden en in welke opzichten zo'n visie op toetsen fundamenteel verschilt van de gangbare opvattingen. Zij laat zien dat met schriftelijke items er meer mogelijkheden zijn wiskundige vaardigheden van kinderen te achterhalen dan tot dusver voor mogelijk werd gehouden.

Van de Rijt beproefde twee versies (een bannende en een sturende) van een remediërend programma met de bedoeling te achterhalen welk van beide versies het meest succes had. Om dit te kunnen onderzoeken ontwikkelde zij toetsen die gebaseerd zijn op de item-response theorie, terwijl zij in inhoudelijke zin uitgaat van de realistische vaktheorie.

Hoewel de drie proefschriften gelet op een aantal theoretisch uitgangspunten grote verwantschap vertonen, hebben we toch te maken met drie studies met een geheel eigen opzet en bedoeling, zoals uit een beknopte schets van de hoofdpunten van de publicaties moge blijken.

Om te beginnen analyseert Van Eerde terugblikkend welke rol de twee funderende wetenschapstradities namelijk de cultuurhistorische school (Vygotskij) en de handelingstheorie (Van

Parreren) enerzijds en de realistische vakdidactische theorie (Freudenthal, Treffers en Goffree) anderzijds in relatie tot elkaar hebben gespeeld in het ontwikkelingsproces. Belangrijke kenmerken van deze theorieën zijn terug te vinden in de opzet van het ontwikkelde instrument: interactie, eigen constructies, open problemen, contexten etc. De auteur laat zien hoe het ontwikkelingsproces verliep, namelijk niet volgens een voorgedachte theorie, maar als een boeiende zoektocht waarbij vele hindernissen genomen moesten worden. In het prille begin werd steun gezocht in het werk van Piaget, maar uit onderzoek bleek dat er weinig samenhang bestond tussen de door Piaget veronderstelde voorwaarden voor getalbegrip (zoals conservatie en klassificatie) en de rekenvaardigheid van de kinderen in de derde groep. Wel wordt – en terecht – gekozen voor Piagets bekende klinische onderzoeksmethode en die keuze wordt uitvoerig verantwoord. Ook de diagnostische werkwijze wordt methodologisch gelegitimeerd en met name aan de interessante kwestie van de waarneembaarheid van handelingen besteedt Van Eerde veel aandacht. Zo is onder meer een verbale handeling waarneembaar, zegt de onderzoekster en hoewel dat waar is, kan het nog wel eens een probleem zijn dat kinderen – als ze hun denken moeten verwoorden – dingen zeggen die mentaal niet hebben plaatsgevonden. Dat dilemma in de diagnostiek is enigszins op te vangen door de kinderen te vragen hard-op te denken. Dat kan echter het oplossingsproces weer verstoren.

Veel zorg besteedde de onderzoekster aan het systematiseren van de diagnostiek, in het bijzonder wat betreft het verzamelen en interpreteren van de gegevens. De auteur maakt helder dat de diagnostische vaardigheid van de gebruiker van het instrument vergroot kan worden door een grondige uitwerking van de observatie- en vraagtechniek. Het project was kortom in handen van een onderzoekster die zich serieus en grondig van haar taak heeft gekwet. Vanzelfsprekend echter roept de tekst hier en daar ook wel vragen op. Bijvoorbeeld de bevinding dat kinderen analoge opgaven soms verschillend oplossen (bijvoorbeeld $8 + 6$ en $9 + 6$). Dit kan als een transferprobleem geïnterpreteerd worden en dat is een serieus maar ook een hardnekkig probleem. Van Eerde laat dit probleem in haar proefschrift onbesproken maar ook de andere

promovendi gaan hier aan voorbij en dat is een beetje jammer.

Inmiddels werd gaandeweg het project enkele externe deskundigen gevraagd om hun oordeel en hun aanbevelingen leidden tot een onderzoek naar de betrouwbaarheid van de kwantijzerwerkwijze en vervolgens tot een nieuwe projectfase waarin het instrument geschikt werd gemaakt voor leerkrachten. De interpretaties van de handelingen van kinderen bleken over het algemeen betrouwbaar, waaruit ook de validiteit van de interpretatietheorie kan worden afgeleid. Aan de ontwikkeling van de Kwantiwijzer voor leerkrachten is veel zorg besteed (nieuwe voorproeven, interpretatieoverzichten etc.) en in de praktijk zal nu nog moeten blijken hoe de leerkrachten met het mooie instrument uit de voeten kunnen.

Laten we deze korte schets van de inhoud van het proefschrift van Van Eerde besluiten met een punt dat wellicht de moeite van een discussie waard is en wel het volgende. Niet beheerste handelingen worden door Van Eerde verklaard uit het niet beheersen van daarvoor benodigde **eenvoudiger** opgaven en handelingen. De vraag luidt of dat altijd wel zo is. Wordt het niet beheersen niet ook (en soms vooral) bepaald door het ontbreken van aan die handelingen ten grondslag liggende **fundamentele inzichten**? Bijvoorbeeld ten aanzien van de opbouw van getallen, de 10- en 5-structuur, basale strategieën? Een hieruit voortvloeiende vraag luidt of in de Kwantiwijzer niet te veel wordt vertrouwd op remediëring via 'lokale' strategieën (zoals overbruggen e.d.)?

In de dissertatie van Van den Heuvel-Panhuizen staat het begrip 'assessment' centraal. 'Assessment', zeg maar het waarderen en interpreteren van de leer- en denkprocessen van kinderen, moet inzicht geven in het wiskundig (leren) denken van kinderen. Expliciet verwijzend naar Vygotskij, meent de auteur dat het er niet om gaat wat een kind **kan**, maar wat het kan **leren**. Het is de bedoeling van deze studie te laten zien wat 'assessment' betekent in het kader van de realistische wiskundige vakdidactiek. De boodschap is helder: als kinderen gestimuleerd worden eigen oplossingen te vinden en geleerde aanpakken creatief toe te passen, moeten ze ook de kans krijgen (als hun kennis wordt getoetst) dat geleerde te laten zien. Dat kan niet op basis van

'gesloten' opgaven, dan moet je ze open context-opgaven voorleggen. Vanuit deze centrale gedachte worden traditionele tests en de funderende taxonomieën (vooral die van Bloom) bekritiseerd (niet valide, te sterk geformaliseerd, denkprocessen blijven verborgen etc.). In het kader van een onderzoek naar de invloed van methoden werd een realistische versie van toets-items ontwikkeld en die ontwikkeling wordt beschreven. Onverwachte ontdekkingen speelden een grote rol onder meer de ontdekking (op basis van interviews) dat kinderen méér konden dan werd verwacht. Dit hing samen met de context. Een verrassing was ook dat 6-jarige kinderen 'open-end' problemen konden oplossen. De kinderen werden gestimuleerd met een kladblaadje te werken en analyse van de aantekeningen leverde boeiende, diagnostische gegevens over het oplossingsproces op.

Wanneer kunnen kinderen echt laten zien wat ze hebben geleerd? Dan moeten de items bestaan uit zorgvuldig gekozen problemen, de problemen moeten betekenisvol zijn, ze moeten een veelheid aan processen uitlokken, ze moeten bestaan uit contexten die toegankelijk zijn en transparant. Ander onderwijs vereist dus andere 'assessment'. De auteur legt de nadruk – en terecht – op de keuze van goede contexten en meent dat kinderen – vooral kinderen uit stimuleringscholen – sterk aan de realiteit vastzitten. Dat kan als daar tenminste mee wordt bedoeld dat bij kinderen met een context een 'discourse' wordt opgeroepen, dat wil zeggen gewoonten, gevoelens, gedrag, rituelen etc. die het denken bepalen. Die 'discourse' kan trouwens ook nog tegengesteld zijn aan de 'discourse' die past bij de wiskunde. En dat is waarschijnlijk een belangrijke reden waarom, zoals de promovenda ook ondervond, niet alle contexten het beoogde handelen opriepen. Inmiddels mag echter vastgesteld worden dat de onderzoekster ons verrast met een mooi onderzoeksresultaat: schriftelijke toetsitems voorgelegd aan nog jonge kinderen leveren veel meer relevante informatie op dan velen tot nu toe dachten. Zij was bovendien vindingrijk in het bedenken van mooie items zoals bijvoorbeeld de 'twin-tasks': op grond van $86 + 57 = 143$ bedenken hoeveel dan $86 + 56$ moet zijn. De auteur maakt velen – ik hoop ook leraren – gevoelig voor hoe (individuele) kinderen denken en leren en hoe je daar achter kunt komen. In het bijzonder zouden leraren in het speciaal onderwijs

kennis moeten kunnen nemen van de wijze waarop het rekenen van hun kinderen werd onderzocht. Overigens zouden, zegt de onderzoekster, deze kinderen ook kunnen reflecteren over hun oplossingswijzen, maar het enige wat ik zag is dat ze hun aanpak noteren of in een tekening weergeven. Als ze dat kunnen, zegt de auteur, 'then they may also be quite capable of talking about it' (p. 253). Misschien is dat zo, maar dan is er nog niet per se sprake van reflectie.

Dat neemt niet weg dat Van den Heuvel-Panhuizen in haar studie op overtuigende manier laat zien hoe een eigen kijk op 'assessment' vanuit een specifieke vaktheorie ontwikkeld kan worden.

Laat ik deze schets met een discussiepunt afsluiten. Ik sluit dan aan bij deze uitspraak van de auteur: 'Large scale assessments.... have been left almost entirely out of the picture' (p. 179). Dat is waar, maar ook een bespreking van of verwijzing naar de met toetsen vaak verbonden methodologische onderbouwing kon ik in het boek niet vinden. Ook niet een verwijzing naar de inmiddels bekende item-response theorie waar het CITO zijn (realistische) toetsen op baseert (Verhelst e.a. 1991). Een verzameling items is echter nog geen toets. Daarmee hangt samen dat het niet gemakkelijk zal zijn bepaalde uitslagen te interpreteren. Wat betekent het als een kind bijvoorbeeld de helft van de voorgelegde items niet aankan? Is dat normaal of moeten er dan interventies overwogen worden? In welk opzicht en in welke mate hebben de items een signalerende functie met het oog op de remediëring?

Zoals dat in de orthopedagogiek vaak gewoon is, verkent Van de Rijt uitvoerig het werk van theoretici die iets belangwekkends over getalbegrip hebben gezegd. Piaget hecht weinig waarde aan het tellen, Freudenthal juist veel en met Mc Shane wordt geconcludeerd dat conservatie als zodanig geen voorwaarde is voor getalbegrip. Er wordt een model van 'rekenvaardigheid' voorgesteld bestaande uit verschillende componenten (zoals vergelijken, seriëren, verkort en resultaatief tellen). Rekenproblemen die zich voordoen in groep 3 zijn terug te voeren tot het onvoldoende beheersen van een of meerdere van deze factoren, zo luidt het centrale uitgangspunt van deze studie. Om de kinderen te helpen, werden er spe-

ciale programma's ontwikkeld waarin de componenten op geïntegreerde wijze werden aangeboden. Er werd geëxperimenteerd met een 'vrijer' programma op basis van 'banende' instructie en een meer 'gebonden' programma met structuur verlenende instructie. Om de effecten van beide programmaversies te kunnen meten en vergelijken, werden drie nieuwe rekenvaardigheidstoetsen ontwikkeld. Het proces van item-constructie (op acht onderdelen) wordt helder verantwoord. Ik moet me steeds weer inspannen om precies te begrijpen wat classificeren met rekenen (telstrategieën) heeft te maken. Ik geloof niet zo veel, al is er natuurlijk wel een verband tussen kardinaal aspect van getal en classificeren. Aan de items werden meetwaarden toegekend en geschaald (gecalibreerd) volgens het item-response-model. Kort gezegd wordt daarbij uitgegaan van de gedachte dat de kans op een juist antwoord met betrekking tot een bepaald item een functie is van de te meten latente vaardigheid van het kind. Op basis van geconstrueerde rekenvaardigheidsschalen werd een longitudinaal onderzoek uitgevoerd waaruit bleek dat de rekenvaardigheid stijgt met de leeftijd. Vervolgens werd een experimenteel programma uitgetoetst met een experimentele (e) en een controle (c) groep en werd een aantal hypothesen getoetst, onder meer de hypothesen dat de e-groep hoger zou scoren dan de c-groep, dat de beide instructievormen evenveel effect sorteren en dat beide meer effect hebben dan methodegebonden onderwijs én meer effect dan de niet methodegebonden aanpak. Inderdaad blijken beide instructievormen vrijwel even effectief. Bij nameting in de derde groep bleek dat de hele e-groep gemiddeld beter scoorde dan de groep die een methode volgde, maar er was geen (significant) verschil met de groep die niet methodegebonden onderwijs kreeg. De vraag welk verband er is tussen voorbereidende rekenvaardigheid en het reken/wiskundeonderwijs in groep drie - aan die vraag wordt niet expliciet aandacht besteed - zou hiermee deels zijn beantwoord. Het experimentele programma geeft immers leerwinst in groep drie. Het is opvallend dat banend onderwijs evenveel effect heeft als 'structuur verlenend' onderwijs; vaak wordt immers gedacht dat een vrijer programma voor kinderen met leerproblemen niet zo geschikt zou zijn. Van de Rijt heeft volgens de regels der kunst haar onderzoekswerk verricht en ze heeft

laten zien dat rekenvaardigheden van jonge kinderen geschaald kunnen worden volgens het item-response model en dat is op zich een verdienste.

Laat ik deze korte weergave van de inhoud weer besluiten met een discussiepunt. Dat betreft de betekenis en de aard van de items waar Van de Rijt voor heeft gekozen. Ten eerste zijn die items - en is de toets - niet erg interactief en wordt de kinderen niet veel ruimte gelaten voor eigen constructies. Maar bovendien zijn de items meestal vrij formeel van aard en dat is niet nodig. Wat zal een kind denken bij de vraag 'Wijsvarken nummer vijf eens aan'? Met weinig moeite hadden de items een aansprekende context kunnen krijgen zodat de opdrachten voor de kinderen als zinvol en leuk worden ervaren. Het is niet ondenkbaar dat dit laatste invloed heeft op de reacties van de kinderen.

De drie dissertatieonderzoeken vertonen op enkele belangrijke punten interessante overeenkomsten maar ook verschillen. Bij alle drie de auteurs vinden we een uitgesproken visie op diagnostiek. Van Eerde benadrukt het belang van interactie en protocolinterpretaties, Van den Heuvel-Panhuizen wil op basis van goede contexten, juiste vragen etc. laten zien dat kinderen méér kunnen dan op basis van traditionele toetsen steeds blijkt en bij Van de Rijt is de signalering gebaseerd op geconstrueerde schaalwaarden. De vakdidactische theorie is bij Van den Heuvel-Panhuizen bepalend voor de visie op diagnostiek, maar speelt bij Van de Rijt meer op de achtergrond een rol. Voor Van Eerde is bovendien de handelingstheorie richtinggevend. Daar staat tegenover dat Van de Rijt het meest expliciet de gehanteerde testtheorie verantwoordt (de item-response theorie) en op basis van de door haar geconstrueerde toetsen zijn tegen de achtergrond van normgroepgegevens in signalerende zin uitspraken mogelijk. In alle drie de dissertaties zijn tenslotte relevante gegevens over rekenzwakke kinderen te vinden en zinvolle aanwijzingen voor hulp aan deze kinderen. Entreetoetsen dus overbodig?

J. M. C. Nelissen

Literatuur

Verhelst, N. D., Verstralen, H., & Glas, K. (1991) *OPLM: The one parameter logistic model and its extension to the polytomous case*. Arnhem: Cito.

Mededelingen

VELON-Congres 1997

Het VELON-Congres 1997, georganiseerd door Stoas en de Landbouwwuniversiteit Wageningen, wordt gehouden op 7 en 8 april 1997 in het Nationaal Sport Centrum Papendal te Arnhem. De titel van het congres is: Studielandschap 2002, een leeromgeving met een goed (voor)uitzicht. Voor nadere informatie: Mevrouw A. van Nieuwenhoven, Stoas, tel.: 0317 47 27 11, fax: 0317 42 47 70.

Promoties

Katholieke Universiteit Brabant

Promovendus: Drs. M. Hajer

Titel proefschrift: *Leren in een Tweede Taal. Interactie in vakonderwijs aan een meertalige MAVO-klas*

Promotor: Prof.Dr. G. Extra

Co-promotor: Dr. C.M.J. van Helvert

Datum: 18 september 1996

Katholieke Universiteit Brabant

Promovendus: Drs. H.A.K. Klatter-Folmer

Titel proefschrift: *Turkse kinderen en hun schoolsucces. Een dieptestudie naar de rol van sociaal-culturele oriëntatie, taalvaardigheid en onderwijskenmerken*

Promotor: Prof.Dr. G. Extra

Co-promotor: Dr. A. Vallen

Datum: 23 september 1996

Rijksuniversiteit Groningen

Promovendus: Drs. M.E. Kalverboer

Titel proefschrift: *Onderzoek naar de wijze van rapporteren over opvoedings- en verzorgingsproblematiek bij de Raad voor de Kinderbescherming directie Noord*

Promotor: Prof.Dr. J.E. Rink

Datum: 3 oktober 1996

Inhoud andere tijdschriften

Comenius

16e jaargang, nr. 3, 1996

Thema: Intercultureel onderwijs

Intercultureel onderwijs en culturele identiteit, door Y. Leeman, H. Lutz en W. Wardekker

'...Van Ausländerpädagogik naar interculturele opvoeding ...', Posing tot een inleidende balans, door M. Krüger-Potratz

Tricks of the trade. Issues in multicultural arts education, door Ph. Cohen
en verder:

'De pijn wakker houden, het verlangen wekken'. Enkele bedenkingen bij de nood aan waardenopvoeding. Deel 1. Karakteropvoeding als alternatief, door J. Masschelein

Jongeren en seksualiteit: inleiding op een artikelserie, door J. Doornenbal en G. Timmermans
Begrensd beleid: overheidsbeleid tegen seksuele intimidatie in het voortgezet onderwijs, door G. Timmermans en Chr. Bajema

Tijdschrift voor Onderwijsresearch

21e jaargang, nr. 3, 1996

Thema: Rendement van het beroepsonderwijs

Rendement van het beroepsonderwijs: Beleid en onderzoek, door W.Th.G.J. Hoebe

Schoolloopbaansucces en de breedte van het opleidingsaanbod in het MBO, door Th.A. van Batenburg en P.R. den Boer

Het intern rendement van het leerlingwezen,

door G.J. Harms en P.R. den Boer

Effecten van de aanwezigheid en inrichting van de praktijkcomponent bij scholing van laag opgeleide volwassenen, door P.R. den Boer

Het duale karakter van beroepsopleidingen en hun waardering door werkgevers, door

W.Th.J.G. Hoebe

en verder:

Impliciet versus expliciet onderwijs in bedrijfs-economische modellen, door A.T.J. Vernooij

Schooladviezen van allochtone leerlingen, door H. Bosma en P. Cremers

Notities Commentaren

Scoringsregels en toetsinstructie, door A. Dirkzwager

Het toekomstperspectief van de onderwijskunde, door L.F.W. de Klerk

Constructieve en destructieve discussies over de interdisciplinaire onderwijskunde. Een reactie op De Klerk, door B. Creemers, N. Lagerweij en N. Verloop

Tijdschrift voor Orthopedagogiek
35e jaargang, nr. 7/8, 1996

Omgaan met verscheidenheid in leren (oratie),
door L. Verhoeven
Commentaar op de oratie, door A. van der Leij
De bestaans(on)zekerheid van pleegkinderen,
door E. Singer
De relatie tussen specifieke taalstoornissen, ge-
drag van het kind en beleving van de ouders,
door S.M. Goorhuis-Brouwer, H. Nakken en
H. van den Berg

Tijdschrift voor Orthopedagogiek
35e jaargang, nr. 9, 1996

Het ethische en de ethiek, door A.B.J. Verstege
Ondersteunende zorg aan gezinnen met een ern-
stig gehandicapt kind, door M.S.M. Waarden-
burg en A. Vermeer
Regionale diagnostiekt teams in de jeugdzorg.
Een inventarisatie van kansen en belemmerin-
gen, door M. Faas

Tijdschrift voor Orthopedagogiek
35e jaargang, nr. 10, 1996

Selectief mutisme: welke vorm van behande-
ling? door N. Cohen de Lara-Kroon
Als mishandelde kinderen ouder worden, door
H.E.M. Baartman
Leerpotentieel: onderzoek bij allochtone leerlin-
gen uit het basisonderwijs, door W.C.M. Resing
en A.M. van Wijk

Ontvangen boeken

- Gent, B. van (red.), *Opleiden en leren in organisaties. Ontwikkelingen, benaderingen, onderzoekin- gen*. Uitgeverij Boom, Meppel, 1996, f 36,50.
- Graaf, G. de, *Mijn kind gaat naar de gewone school. Integratie van kinderen met een verstandelijke belemmering*. Acco, Leuven/Amersfoort, 1996, f 46,75.
- Harpman, A (red.), *Stimulerend opvoeden. Ontwik- kelingsstimulering van kinderen van 0 tot 4 jaar, opvoedersamenwerking en het aanleren van pedagogische vaardigheden*. Samsom H.D. Tjeenk Willink, Alphen aan den Rijn, 1996, f 39,75.
- Harpman, A., & Tavecchio, L. (red.), *Opvoeding als evenwichtskunst. Pedagogiek voor 0 tot 2 jari- gen*. Samsom H.D. Tjeenk Willink, Alphen aan den Rijn, 1996, f 47,50.
- Hopman, M. (red.), *Opvoeden in kindercentra. Visies, wetenschappelijke ontwikkelingen en praktijk*. De Tijdstroom, Utrecht, 1996, f 69,-.
- Klaassen, C., *Socialisatie en moraal. Onderwijs en waarden in een laat-moderne tijd*. Uitgeverij Ga- rant, Leuven/Apeldoorn, 1996, f 39,50.
- Rietdijk, L.E., *Behandeling in de leefgroep; naar een werkmodel*. SWP, Utrecht, 1996, f 52,50.
- Veugelers, W., & Zijlstra, H. (red.), *Netwerken aan de bovenbouw van HAVO en VWO. Scholen en na- scholing op weg naar het studiehuis*. Uitgeverij Garant, Leuven/Apeldoorn, 1995, f 39,50.
- Vlaskamp, C., Blokhuis, A., & Ploemen, M., *Gewoon bijzonder. Opvoeden van kinderen met een ern- stige meervoudige handicap*. Van Gorcum, As- sen, 1996, f 25,-.
- Vygotskij, L., *Cultuur en ontwikkeling*. Inleiding, ver- taling en redactie: R. van der Veer. Boom, Mep- pel, 1996.

C.A.J. Aarnoutse (K.U.Nijmegen)
J.J. Beishuizen (R.U.Leiden)
R. van den Berg (K.U.Nijmegen)
Th. Bergen (K.U.Nijmegen)
P. De Boeck (K.U.Leuven)
R.J. Bosker (U.T.Enschede)
K. Bügel (Cito, Arnhem)
A.G. Bus (R.U.Leiden)
B.P.M. Creemers (R.U.Groningen)
J. Elen (K.U.Leuven)
W. van Esch (I.T.S.Nijmegen)
P. Gesquière (K.U.Leuven)
E.G. Harskamp (R.U.Groningen)
J. Hendriksen (Inst.Kindervalidatie, Valkenburg)
W.Th.J.G. Hoebe (R.U.Groningen)
R.H. Hofman (R.U.Groningen)
W.K.B. Hofstee (R.U.Groningen)
J. van Hout (U.v.A.Amsterdam)

B.H.A.M. van Hout-Wolters (U.v.A.Amsterdam)
P.J. Janssen (K.U.Leuven)
J.J. van Kuyk (Cito, Arnhem)
A. van der Leij (U.v.A.Amsterdam)
P.P.M. Leseman (U.v.A.Amsterdam)
E.C.D.M. van Lieshout (K.U.Nijmegen)
C.J.W. Meijer (R.U.Groningen)
G. Schellings (K.U.B.Tilburg)
K. Sijtsma (U.U.Utrecht)
J. Terwel (U.v.A.Amsterdam)
A. Treffers (U.U.Utrecht)
T. Vallen (K.U.B.Tilburg)
S.A.M. Veenman (K.U.Nijmegen)
L. Verhoeven (K.U.Nijmegen)
J. Vermunt (R.U.Leiden)
A.M.L. van Wieringen (U.v.A.Amsterdam)
R. Wierstra (U.U.Utrecht)
Th. Wubbels (U.U.Utrecht)