Utrecht University Library
Partner in Science
Universiteit Utrecht

Pedagogische studiën 1993 : Tijdschrift voor onderwijskunde en opvoedkunde

https://hdl.handle.net/1874/205488

So. oct. 6441

Pedagogische Studiën

Tijdschrift voor onderwijskunde en opvoedkunde

Inhoud Jaargang 70 1993

RIJKSUNIVERSITEIT TE UTRECHT
2518 181 3

Bijlage bij *Pedagogische Studiën*, nr. 6, 1993

Pedagogische Studiën

Tijdschrift voor onderwijskunde
en opvoedkunde

Redactie

Prof. Dr. W. J. Nijhof (voorzitter)
Dr. P. J. J. Stijnen (secretaris)
Prof. Dr. L. J. Th. van der Kamp
Prof. Dr. A. van der Leij
Prof. Dr. C. F. M. van Lieshout
Prof. Dr. G. W. Meijnen
Prof. Dr. P. R. J. Simons
Dr. J. H. Slavenburg
Prof. Dr. N. Verloop
Prof. Dr. L. Verschaffel

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

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

Inhoudsopgave

Jaargang 70, 1993

Auteursregister

Artikelen

Amelsvoort, H. W. C. H. van, Bosker, R. J., & Brandsma, T. F.
Rendementsbekostiging in de bovenbouw van HAVO en VWO; overwegingen en reacties 283

Bon, W. H. J. van
zie: Theloosen, G.

Bosker, R.J.
zie: Amelsvoort, H. W. C. H. van

Bosman, A. M. T. M., & Leerdam, M. van
Aanvankelijk spellen: de dominatie van de verklankende spelwijze en de geringe effectiviteit van lezen als spellingsinstructie en -methode 28

Brandsma, T. F.
zie: Amelsvoort, H. W. C. H. van

Damhuis, R., & Glopper, K. de
Leerkrachtoordelen over de wenselijke woordenschhat van 4- tot 6-jarigen: betrouwbaarheid, validiteit en gebruiksmogelijkheden 322

Dekkers, H.
Determinanten van seksespecifieke vakkenkeuzen in HAVO en VWO 445

Derriks, M., & Kat. E. de
Onedele motieven voor een edel beroep? Een studie naar de motieven om te kiezen voor een hogere pedagogische opleiding 17

Driessen, G.
Binnen en buiten Onderwijsvoorrangsgebieden. Sociaal-etnische achtergronden en onderwijspositie van leerlingen in het voortgezet onderwijs 252

Glopper, K. de
zie: Damhuis, R.

Harskamp, E. G.
Spelen op school 337

Hoeksma, J. B.
zie: Koomen, H. M. Y.

Kappers, E.J.
zie: Walraven, A. M. A.

Kat, E. de
zie Derriks, M.

Kieviet, F. K.
Visitaties lerarenopleidingen nader bekeken 274

Kloprogge, J.
zie: Meijnen, G. W.

Koomen, H. M. Y., Hoeksma, J. B., & Meerum Terwogt, M.
Het belang van informatie over een afwezige genetische vader 368

Lagerweij, N. A. J.
Inplementatie van basisvorming: de start van een dieptestudie 162

Leenders, Y.
zie: Veenman, S.

Leerdam, M. van
zie: Bosman, A. M. T. M.

Leij, A. van der
zie: Smeets, H.
zie: Yap, R.

Meerum Terwogt, M.
zie: Koomen, H. M. Y.

Meijnen, G. W., & Kloprogge, J.
De evaluatie van het Onderwijsvoorrangsbeleid. Een introductie op de artikelen 82

Meijnen, G. W., & Kloprogge, J.
De evaluatie van het Onderwijsvoorrangsbeleid geëvalueerd 84

Mey, J. R. P. B. de
Morele oordeelsvorming bij 'antisociale' jongeren en scholieren 46

Meyer, P.
zie: Veenman, S.

Mulder, L.
De invloed van het advies op de schoolloopbaan van de OVB-doelgroepleerlingen in het voortgezet onderwijs 242

Mulder, L.
zie: Werf, M. P. C. van der

Perrenet, J. Chr.
De invloed van het gebruikte leerboek op het maken van specifieke fouten bij wiskundeproblemen 195

Raemaekers, J.
zie: Roelofs, E.

Reitsma, P.
zie: Sliepen, S. E.
zie: Walraven, A. M. A.

Roelofs, E., Veenman, S., & Raemaekers, J.
Teamgerichte nascholing en coaching: Effecten op het didactisch handelen van leerkrachten in combinatieklassen 352

Sanders, M.
zie: Veenman, S.

Sliepen, S. E., & Reitsma, P.
Training van lom- leerkrachten in directe instructie van begrijpend leesstrategieën 420

Smeets, H., & Leij, A. van der
Differentiële effecten van computergestuurde instructie op het lezen van dyslectische en zwak lezende kinderen: een voorstudie 56

Stokking, K. M.
Differentiatie en zorgverbreding in het basisonderwijs: de behandeling van een vernieuwingsonderwerp in onderzoek en beleid, 1984-1990 135

Tesser, P., & Vierke, H.
De prestaties van de OVB-doelgroepen in 1990; stand van zaken na de tweede meting 122

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

Theloosen, G., & Bon, W. H. J. van
Herhaling van woorden en lettergroepen: oefeneffecten bij zwakke lezers 180

Veenman, S., Leenders, Y., Meyer, P., & Sanders, M.
Leren lesgeven met het directe instructiemodel 2

Veenman, S.
zie: Roelofs, E.

Verhaak, C.
Samenwerking in Onderwijsvoorrangsgebieden 263

Vierke, H.
zie: Tesser, P.

Walraven, A. M. A., Reitsma, P., & Kappers, E. J.
Instructie in strategieën voor begrijpend lezen aan zwakke lezers 298

Weide, M. G.
zie: Werf, M. P. G.

Werf, M. P. C. van der, Tesser, P., & Mulder, L.
De prestaties van de OVB-doelgroepen in 1988; stand van zaken na de eerste meting 97

Werf, M. P. C. van der, & Weide, M. G.
Effectieve voorrangskenmerken in de school en in de klas 108

Yap, R., & Leij, A. van der
Computergestuurde remediëring van dyslexie door het opvoeren van de herkenningssnelheid van subwoordeenheden 402

Discussiebijdragen

Baggen, P.
zie: Weijers, I.

Knoers, A. M. P.
De autonomievergroting van scholen. Zin en onzin 382

Masschelein, J.
De dood van de universiteit 206

Masschelein, J.
Het belang van onderscheid 217

Weijers, I., & Baggen, P.
'Storen' - De laatste strohalm van het idealisme 214

Kronieken

Imants, J.
De jaarlijkse bijeenkomst van het 'International Congress for School Effectiveness and Improvement': Norköping 1993 220

Slavenburg, J. H.
Voorschoolse stimuleringsprojecten 309

Stijnen, P. J. J. (red.)
De Onderwijsresearchdagen 1993 in Maastricht: 26-28 mei, met bijdragen van *J. J. H. van den Akker, R. Bosker, G. ten Dam, J. van der Linden, G. Rijlaarsdam, P. Sanders, J. van der Sanden, P. J. J. Stijnen, J. Terwel, M. G. Weide, J. Willems* 387

Verloop, N. (red.)
Het AERA-Congres 1993, Atlanta, met bijdragen van *J. J. H. van den Akker, B. P. M. Creemers, S. Dijkstra, D. Dolmans, T. de Jong, J. J. G. Merriënboer, P. Sanders, N. Verloop* 460

Boekbesprekingen

Bosch, K. van den
Poor readers' decoding skills. Effects of training, task, and word characteristics (*R. Yap*) 151

Freudenthal, H.
Revisiting mathematics education. Chinese lectures (*J. M. C. Nelissen*) 74

Goffree, F.
Wiskunde en Didactiek 2 (*H. Stroomberg*) 154

Hofman, R. H.
Effectief schoolbestuur. Een studie naar de bijdrage van schoolbesturen aan de effectiviteit van basisscholen (*A. M. den Haan*) 226

Meijnen, G. W., & Riemersma, F. S. J.
Schoolcarrières: een klassenkwestie? De schoolloopbanen van leerlingen, bezien in relatie tot de invloeden die maatschappelijke en binnenschoolse determinanten hierop hebben (*P. Jungbluth*) 158

Peer, D. R. van
Sociale ondersteuning van kinderen met leerstoornissen (*J. Th. A. Bakker*) 155

Peetsma, T. T. D.
Toekomst als motor? Toekomstperspectieven van leerlingen in het voortgezet onderwijs en hun inzet voor school (*M. J. de Jong*) 76

Ruijssenaars, A. J. J. M.
Rekenproblemen. Theorie, diagnostiek, behandeling (*J. M. C. Nelissen*) 475

Sijtstra, J. M.
Doel en inhoud van het taalonderwijs. De ontwikkeling van een model voor domeinbeschrijvingen van taalonderwijs (*G. Rijlaarsdam*) 229

Tulder, M. van
Nascholing en onderwijsvernieuwing (*J. H. Slavenburg*) 316

Vreeke, G. J.
Zorg en rechtvaardigheid. Analyse van de Kohlberg-Gilligandiscussie (*A. W. van Haaften*) 73

Vreugdenhil, C.
De Führungslehre van Peter Petersen. Historisch-pedagogisch tekst- en contextonderzoek van de onderwijspedagogiek van het Jenaplan (*J. D. Imelman*) 222

Thema: Evaluatie
Onderwijsvoorrangsbeleid

Nummer 2

G. W. Meijnen & J. Kloprogge
De evaluatie van het Onderwijsvoorrangsbeleid. Een introductie op de artikelen 82

G. W. Meijnen & J. Kloprogge
De evaluatie van het onderwijsvoorrangsbeleid geëvalueerd 84

M. P. C. van der Werf, P. Tesser & L. Mulder
De prestaties van de OBV-doelgroepen in 1988: stand van zaken na de eerste meting 97

M. P. C. van der Werf & M. G. Weide
Effectieve voorrangskenmerken in de school en in de klas 108

P. Tesser & H. Vierke
De presentaties van de OVB-doelgroepen in 1990: stand van zaken na de tweede meting 122

Nummer 4

L. Mulder
De invloed van het advies op de schoolloopbaan van OVB-doelgroepleerlingen in het voortgezet onderwijs 242

G. Driessen
Binnen en buiten Onderwijsvoorrangsgebieden. Sociaal-etnische achtergronden en onderwijspositie van leerlingen in het voortgezet onderwijs 252

C. Verhaak
Samenwerking in Onderwijsvoorrangsgebieden 263

Inhoudsopgave

Pedagogische Studiën Jaargang 70, nummer 1/1993

S. Veenman, Y. Leenders, P. Meyer en M. Sanders
Leren lesgeven met het directe instructiemodel 2

M. Derriks en E. de Kat
Onedele motieven voor een edel beroep?
Een studie naar de motieven om te kiezen voor een hogere pedagogische opleiding 17

A. M. T. M. Bosman en M. van Leerdam
Aanvankelijk spellen: de dominantie van de verklankende spelwijze en de geringe effectiviteit van lezen als spellinginstructiemethode 28

J. R. P. B. de Mey
Morele oordeelsvorming bij 'antisociale' jongeren en scholieren 46

H. Smeets en A. van der Leij
Differentiële effecten van computergestuurde instructie op het lezen van dyslectische en zwak lezende kinderen: een voorstudie 56

Boekbesprekingen 73

Mededelingen 78

BIBLIOTHEEK DER
RIJKSUNIVERSITEIT
UTRECHT

Leren lesgeven met het directe instructiemodel*

S. Veenman, Y. Leenders, P. Meyer en M. Sanders**

Samenvatting

Een quasi-experimenteel onderzoek (experimentele – controlegroep) is uitgevoerd om de effecten te meten van een voor de lerarenopleiding basisonderwijs (PABO) ontwikkeld trainingsprogramma over effectieve instructie. In dit programma zijn twee modellen voor directe instructie uitgewerkt: een model voor expliciete of uitvoeringsvaardigheden en een model voor impliciete of hogere cognitieve vaardigheden. Deze modellen baseren zich op de uitkomsten van onderzoek over effectieve instructie. Voorafgaande en na afloop van het trainingsprogramma zijn bij PABO-studenten observaties uitgevoerd door getrainde observatoren en door de mentoren van deze studenten. Ook is het taakgerichte gedrag van de leerlingen vastgesteld waaraan de studenten op de oefenschool hebben lesgegeven. Verder zijn evaluatiegegevens verzameld over de inhoud en de opzet van het trainingsprogramma. Uit de gegevens blijkt dat de studenten na het volgen van het trainingsprogramma de geselecteerde instructiegedragingen beter uitvoeren dan vóór het trainingsprogramma. Het programma wordt door PABO-studenten en -docenten positief beoordeeld.

Inleiding

Het effect van onderwijs is grotendeels afhankelijk van de kwaliteit van de instructie. Gelet op het streven het beroep van leraar verder te professionaliseren, is het wenselijk dat de opleiding van leraren zich op de hoogte stelt van de nieuwste kennis en inzichten over het geven van effectieve instructie en haar studenten leert hun handelen daarmee in overeenstemming te brengen. "Een professioneel werkend opleidings- en nascholingsinstituut zal zich op de hoogte houden van ontwikkelingen in wetenschappelijk onderzoek en theorievorming in uiteenlopende richtingen, zoals de orthopedagogiek en orthodidactiek en op het terrein van vakinhouden, vakdidactieken en de onderwijswetenschap" (ARBO, 1988, p. 29). Onderzoek heeft de laatste jaren een groot aantal aanwijzingen opgeleverd die bij toepassing ervan leiden tot effectieve instructie (Creemers, 1991; Veenman, 1992). Als maat voor effectiviteit geldt hierbij de leerprestaties van de leerlingen. Hoewel onderzoek een groot aantal kenmerken met betrekking tot effectieve instructie heeft vastgesteld, wordt dit kennisbestand bij het leren lesgeven nog weinig gebruikt. Vanuit de gedachten dat het belangrijk is dat aanstaande leraren getraind worden in vaardigheden die kunnen leiden tot effectieve instructie en dat de opleiding tot leraar in het basisonderwijs de laatste jaren voorwerp is geweest van openlijke kritiek over haar kwaliteit (zie o.a. ARBO, 1988; Ministerie van Onderwijs en Wetenschappen, 1990), hebben het Christelijk Pedagogisch Studiecentrum (CPS) en de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen een trainingsprogramma of module ontwikkeld over effectieve instructie. Bij de ontwikkeling hiervan is gebruik gemaakt van recente onderzoeksinzichten over effectief instructiegedrag. Hierbij speelde ook de gedachte mee dat het model van Didactische Analyse van Van Gelder voor veel opleidingen een aantal tekorten kent. Het model is te veel prescriptief van aard. "Het weerspiegelt niet de feitelijkheid van het onderwijs, maar het geeft weer hoe de onderwijs zou moeten plaatsvinden" (Creemers,

* Dit onderzoek werd mede mogelijk gemaakt door het Christelijke Pedagogisch Studiecentrum (CPS), Hoevelaken.

** De auteurs bedanken de volgende personen voor hun bijdrage aan deze studie: K. Vernooy (CPS), G. van Hardeveld (CPS), M. Förrer (CPS), H. Harings (RTD KUN), M. Voeten (Vakgroep Onderwijskunde KUN); en de deelnemende PABO-docenten, mentoren en studenten voor hun medewerking.

1991, p. 11). Verder geeft het model weinig aanknopingspunten voor de uitvoering van het instructieproces.

De nieuw ontwikkelde module is gedeeltelijk gebaseerd op de positieve bevindingen omtrent het teamgerichte nascholingsprogramma 'Omgaan met Combinatieklassen'(Veenman, Lem & Nijssen, 1988; Veenman, Lem, Roelofs & Nijssen, 1992; Veenman, Lem & Roelofs, 1990). Het onderdeel over effectieve instructie uit dit programma vormde het aangrijpingspunt voor een bewerking voor de lerarenopleiding basisonderwijs (PABO).

1 Directe instructie

Bij de uitwerking van de module over effectieve instructie is gekozen voor het directe instructiemodel. Dit model wordt zo genoemd, omdat de uitvoering van het onderwijsleerproces, zeker in de beginfase, gecontroleerd wordt door de leraar. Het directe instructiemodel met zijn systematische methoden voor het aanbieden van nieuwe leerstof in kleine stappen, het controleren of de leerlingen de leerstof begrepen hebben en het stimuleren tot actieve en succesvolle deelname aan het onderwijsgebeuren, is in het basis-, speciaal en voortgezet onderwijs bij lezen en rekenen/wiskunde zeer succesvol gebleken. Maar ook bij vakken of leerstofonderdelen als vreemde talen, grammatica, natuurkunde, scheikunde, biologie, algebra en onderdelen van wereldoriëntatie (Bickel & Bickel, 1986; Brophy & Good, 1986; Mastropieri & Scruggs, 1987; Rosenshine & Stevens, 1986; Walberg, 1991). Bovendien blijkt het model zeer effectief te zijn voor kinderen uit achterstandssituaties (Carnine, 1989). Deze kinderen die meestal niet beschikken over zelfregulerende studievaardigheden of door een handicap weinig succeservaringen kennen op school, hebben veel behoefte aan actieve instructie en begeleiding door de leraar (Brophy, 1991).

Hoewel het begrip directe instructie meerdere betekenissen heeft, wordt het in deze studie gebruikt als een samenstel van instructiegedragingen die volgens de synthese van Rosenshine en Stevens (1986) in meerdere correlationele en experimentele studies effectief is gebleken. De kern van dit samenstel van effectief gebleken gedragingen bestaat uit de volgende hoofdfasen: a) terugblik op en activering van voorafgaande leerstof; b) presentatie en uitleg van de nieuw te leren leerstof met inbegrip van laten zien/voordoen; c) nauwgezette inoefening en begeleiding: in deze fase oefenen de leerlingen wat ze zojuist geleerd hebben en krijgen hierbij corrigerende feedback van de leraar; d) zelfstandige of individuele oefening: in deze fase zetten de leerlingen de stap van het integreren van de nieuwe kennis of vaardigheid met de reeds aanwezige kennis naar de fase van de automatisering en e) periodieke herhaling van de leerstof. Dit instructiemodel is vooral geschikt voor leerstof en vaardigheden waarbij alle stappen die een leerling moet doorlopen om iets te leren vooraf geëxpliciteerd kunnen worden en afzonderlijk kunnen worden geoefend. De leerstof die hierbij gebruikt wordt laat zich meestal goed structureren en opdelen in kleine stappen. Het handelen kan regelrecht op zijn doel afstevenen, de leertaken kunnen volgens een vast schema worden afgewikkeld. Van Parreren (1988) spreekt in dit verband van uitvoerend handelen. Rosenshine en Edmonds (1990) spreken van 'expliciete kennis en vaardigheden'. Met name is het geschikt voor het aanleren van rekenkundige algoritmen en wiskundige berekeningen, spellingsregels, leren decoderen, het hanteren van een spoorboekje, het lezen van een wegenkaart, grammaticale regels bij vreemde talen, het hanteren van werktekeningen bij het repareren van een auto, het leren omgaan met een tekstverwerkingspakket op de computer. In deze studie wordt dit instructiemodel het directe instructiemodel voor *expliciete vaardigheden* genoemd. Dit model is sterk geïnspireerd door behaviouristische en taak-analytisch georiënteerde opvattingen over leren. Aangenomen wordt dat vaardigheden kunnen worden uiteengelegd in deelvaardigheden die afzonderlijk kunnen worden geleerd.

Het directe instructiemodel voor expliciete vaardigheden is minder geschikt voor vaardigheden of leerstofgebieden die zich minder makkelijk laten structureren, waar de te onderwijzen kennis of vaardigheden niet stapsgewijs aangeleerd kunnen worden. Voorbeelden hiervan zijn begrijpend lezen, het schrijven van een opstel, het oplossen van wiskundige problemen, het begrijpen van gedichten of historische

gebeurtenissen. Hierbij gaat het om complexe cognitieve vaardigheden. In het recente onderzoek naar effectieve instructie, dat gebaseerd is op de cognitieve psychologie, krijgen instructievormen die gericht zijn op het ontwikkelen van deze hogere cognitieve vaardigheden veel aandacht. In deze cognitief georiënteerde visie op instructie wordt een belangrijke plaats ingeruimd voor het strategisch handelen. In tegenstelling tot uitvoerend handelen dat gericht is op het bereiken van een eindresultaat, ligt het doel van strategisch handelen in volgende, nieuwe handelingen. 'Men handelt strategisch om uitvoerende handelingen beter te laten verlopen' (Van Parreren, 1988). Hierbij spelen metacognitieve vaardigheden een belangrijke rol. "Metacognitie omvat de concrete, waarneembare, cognitieve activiteiten waarmee leerlingen zich oriënteren op een leertaak, hun leerproces bewaken ('monitoring'), toetsen, diagnostiseren en sturen tijdens het uitvoeren van een leertaak" (De Jong, 1992, p. 14). Als een leerling zich bewust is van wat hij over een bepaald onderwerp weet en kan (en wat niet), is de kans groter dat hij op het juiste moment de juiste kennis en vaardigheden weet in te zetten (Prawat, 1989; Simons, 1990). In de terminologie van Rosenshine en Edmonds (1990) is dit strategisch handelen gericht op wat zij noemen 'impliciete vaardigheden', vaardigheden waarbij het handelen niet altijd regelrecht op zijn doel kan afstevenen, waarbij de leertaak niet volgens een vast schema kan worden afgewikkeld. In hun meest recente overzicht van effectief instructiegedrag spreken Rosenshine en Meister (1992) niet meer van 'impliciete vaardigheden' maar van 'higher-level cognitive strategies', waarbij strategieën verwijzen naar manieren van aanpak. Het geven van expliciete instructie in het ontwikkelen en toepassen van hogere cognitieve denkprocessen dat gericht is op het leren leren van de leerling steunt op de uitkomsten van het onderzoek van de cognitieve strategie-instructie (Pressley, Goodchild, Fleet & Zajchowski, 1989) en van het 'cognitief leerling-gezel-meester-systeem' (Collins, Brown, & Newman, 1989; Resnick, 1987).

Een belangrijk organiserend concept in het onderwijzen van hogere cognitieve kennis en vaardigheden is het begrip '*scaffolding*': van steigers voorzien, schragen, ondersteunen. Dit begrip is nauw verwant aan het begrip 'zone van naaste ontwikkeling' van Vygotsky (1978). Hiermee wordt bedoeld de steun die de leerling van de leraar krijgt bij het volbrengen van taken die zonder die hulp te moeilijk zouden zijn. De leraar reduceert de complexiteit van de situatie, geeft structuur, verheldert het probleem, wijst de leerling op de volgende stap die gezet dient te worden, bewaakt het doel van de activiteit, betrekt de leerling bij het volbrengen van een gezamenlijke taak, schept een kader en geeft regels die de leerling geleidelijk aan overneemt (zelfregulatie). Zodra de leerling de taak zelfstandig kan volbrengen, wordt de ondersteuning beëindigd, zoals een steiger wordt weggehaald als die niet meer nodig is. Voorbeelden van steigers of ondersteuning zijn: voordoen of modelleren (in de zin van: model staan voor), hardop denken als expert waardoor mentale processen 'zichtbaar' en toegankelijk worden gemaakt (cf. Schoenfeld, 1985), het geven van procedurele steun (cf. Bereiter & Scardamalia, 1987), rolwisselend of wederzijds onderwijzen (cf. Palincsar & Brown, 1984), aanwijzingen en ondersteuning door de leraar als coach, het laten zien van voorbeeldig uitgewerkte opdrachten waarmee de leerlingen hun werk kunnen vergelijken, het geven van lijsten met kritische aandachtspunten voor het maken van een leertaak, leren van en aan elkaar in kleine groepen. Hoewel het begrip 'scaffold' ook bruikbaar is voor directe instructie in expliciete vaardigheden, is de uitwerking het grootst bij directe instructie voor impliciete of hogere cognitieve vaardigheden (Rosenshine & Edmonds, 1990; Rosenshine & Meister, 1992).

Het basisstramien van het instructiemodel voor impliciete of hogere cognitieve vaardigheden is grotendeels gelijk aan dat van het instructiemodel voor expliciete vaardigheden: terugblik en activering van reeds behandelde leerstof, expliciete instructie (in de cognitieve activiteit), het geven van nauwgezette begeleiding, zelfstandige verwerking van de leerstof, feedback en het geven van correcties. De invulling van dit basisstramien is echter verschillend. Er wordt meer gebruik gemaakt van instructievormen die gericht zijn op zelfregulatie, op het stimuleren van de reflectie van de leerlingen op hun eigen cognitieve vaardigheden en op het bevorderen van autonoom leren.

Geïnspireerd door het werk van Rosenshine

en Edmonds (1990) is in de module voor de PABO het basisstramien van het directe instructiemodel voor expliciete vaardigheden aangevuld met instructievormen die afkomstig zijn uit de cognitieve strategiebenadering. Het onderdeel 'presentatie' is in de module aangevuld met instructiegedragingen als: het gebruik maken van handelingswijzers of stappenplannen, het hardop denkend de vaardigheid voordoen (modelleren). Het onderdeel 'het geven van nauwgezette begeleiding' is uitgebreid met instructiegedragingen als: het hanteren van handelingswijzers, het werken met half afgemaakte voorbeelden, het geleidelijk verhogen van de complexiteit van de leertaak, rolwisselend of wederzijds onderwijzen, het laten werken van leerlingen in kleine groepen. Het onderdeel 'individuele verwerking' is aangevuld met instructiegedragingen als het geleidelijk terugtrekken van de gegeven ondersteuning en het leren toepassen van het geleerde in verschillende gebruikscontexten. Deze aanvulling van het instructiemodel wordt in de PABO-module het directe instructiemodel voor *impliciete vaardigheden* genoemd. Beide modellen worden in de module uitgelegd en toegelicht. Via een aantal uitgewerkte lesvoorbeelden worden de studenten gewezen op de mogelijke gebruiksvormen van beide instructiemodellen. Gekozen is voor een behandeling van beide instructiemodellen in één module omdat instructiegedragingen als modelleren, hardop denken, het gebruik van handelingswijzers en het laten werken van de leerlingen in kleine groepen ook toegepast kunnen worden in het instructiemodel voor expliciete vaardigheden. "This suggests that instead of a dichotomy, there is a continuum from well-structured explicit skills to cognitive strategies" (Rosenshine & Meister, 1992, p. 32). Naarmate het accent meer komt te liggen op cognitieve strategieën in plaats van op expliciete vaardigheden neemt de belangrijkheid van het gebruik van 'scaffolds' toe.

2 Onderzoeksvragen

Deze studie wil het uit onderzoek succesvol gebleken onderwijsgedrag met betrekking tot directe instructie vertalen naar een trainingsprogramma of module voor de opleiding van leraren basisonderwijs (PABO) en evalueren op implementatie. Deze evaluatie kent de volgende vraagstellingen die gespecificeerd zijn naar drie niveaus van effecten:

a) het reactieniveau: hoe hebben de studenten en docenten van de PABO de module ervaren?

b) het gedragsniveau: passen de studenten de in de module uitgewerkte effectieve instructiegedragingen toe in de lessen die zij op de oefenschool geven?

c) het leerlingniveau: heeft implementatie van de gekozen effectieve instructiegedragingen invloed op het taakgericht gedrag van de leerlingen van de oefenschool?

3 Opzet van het onderzoek

Het onderzoek bestaat uit drie deelonderzoeken:

1) een deelstudie met interviews en vragenlijsten dat gericht is op de waardering van het trainingsprogramma door de studenten (n=130) en de docenten (n=4) (vraagstelling 1);
2) een observatie-onderzoek waarbij gebruik wordt gemaakt van getrainde observatoren en dat gericht is op de toepassing van de kennis en vaardigheden uit het trainingsprogramma door de studenten op de oefenschool en op de verhoging van het taakgerichte gedrag van de leerlingen. Dit observatie-onderzoek is opgezet volgens een quasi-experimenteel 'pretest-posttest control group design' met een experimentele groep (n=27) en een controlegroep (n=14) (vraagstelling 2 en 3);
3) een observatie-onderzoek waarbij gebruik wordt gemaakt van de beoordelingen van de mentoren (n=41) van de studenten en dat gericht is op de toepassing van de kennis en vaardigheden uit het trainingsprogramma door de studenten op de oefenschool. Deze deelstudie met observaties is opgezet volgens een 'then-post design', een aanpassing van het 'pre, then, and post' design (Mezoff, 1981) (vraagstelling 2).

4 Methode en instrumenten

4.1 De onderzoeksgroep

De onderzoeksgroep bestaat uit tweedejaars studenten van drie PABO's. De PABO's op locaties A en B zijn gelegen in het midden-oosten van het land, de PABO op locatie C ligt in het noord-westen van het land en heeft twee vestingen in verschillende plaatsen. Van locatie A zijn vier klassen geselecteerd voor deelname aan het trainingsprogramma 'Effectieve Instructie', van locatie B één klas en voor locatie C drie klassen. De studenten uit deze 8 klassen (n=205) hebben het programma met betrekking tot effectieve instructie gevolgd (experimentele groep). Op locatie A en B is het programma gegeven door één docent, op locatie C door twee docenten. Op elke locatie hebben studenten uit parallelklassen (n=145) het trainingsprogramma niet gevolgd (controlegroep). Alle deelnemende studenten hebben tijdens de uitvoering van het programma één dag in de week stage gelopen op een basisschool.

De vragenlijsten die gebruikt zijn voor het verzamelen van informatie over de waardering van het trainingsprogramma en het gebruik van de instructiemodellen in de klas, zijn toegestuurd aan alle studenten uit de experimentele klassen (n=205). Van deze studenten hebben 130 de vragenlijst geretourneerd (respons = 63%). De interviews zijn afgenomen bij 4 experimentele klassen (op locatie A en B bij één klas, op locatie C bij twee klassen) en 4 docenten.

Voor het observatie-onderzoek met de getrainde observatoren is om praktische redenen (geld en menskracht) uit de totale experimentele groep op aselecte wijze een beperkt aantal studenten gekozen voor de observaties. Uit de experimentele klassen zijn 27 studenten van locatie A en B geselecteerd (18 van locatie A en 9 van locatie B). Uit de controlegroep zijn 14 studenten geselecteerd voor de observaties (7 van locatie A en 7 van locatie B). Ook van locatie C zijn oorspronkelijk studenten geselecteerd voor deelname aan het observatie-onderzoek (18 voor de experimentele groep en 9 voor de controlegroep). Vanwege een misverstand hebben de docenten aan de observatoren niet doorgegeven wanneer zij het trainingsprogramma zouden geven. Op het moment dat de observatoren de voormeting op de locaties A en B hadden afgerond en de docenten van locatie C benaderden voor afspraken omtrent de observaties, bleek het trainingsprogramma reeds gegeven te zijn. Doordat van deze locatie geen informatie met betrekking tot de voormeting verkregen kon worden, is besloten bij deze locatie geen observaties voor de nameting te verrichten.

Na afloop van het trainingsprogramma is aan de mentoren gevraagd de verrichtingen van hun studenten aan de hand van hun observaties te beoordelen. Van de 150 mentoren die de beoordelingsschaal toegestuurd kregen, hebben slechts 61 de beoordelingsschaal met enkele aanvullende vragen teruggestuurd (respons = 41%). Van deze 61 geretourneerde beoordelingsschalen zijn er 41 bruikbaar gebleken voor verdere analyse (totale respons = 27%).

4.2 De Directe Instructie Schaal

Zowel de getrainde observatoren als de mentoren van de oefenschool hebben gebruik gemaakt van de Directe Instructie Schaal (DIS) om het instructiegedrag van de studenten te beoordelen. Deze beoordelingen zijn gegeven aan de hand van een vierpuntsschaal die aangaf in hoeverre studenten gebruik hebben gemaakt van de instructiegedragingen uit het instructiemodel voor expliciete en impliciete vaardigheden. Beoordeeld is het algemeen gebruik van de instructiegedragingen in de context van twee soorten lesopdrachten (zie hieronder). De 20 items uit de DIS zijn gebaseerd op het werk van Rosenshine en Stevens (1986), Rosenshine en Edmonds (1990) en op de Management en Instructie Schaal (MIS) die gebruikt is in ons eerder onderzoek (Veenman, Lem & Roelofs, 1990; Roelofs, Raemaekers & Veenman, 1991). Voor de hanteerbaarheid van het instrument zijn in de DIS niet alle instructiegedragingen uit het trainingsprogramma opgenomen. Gekozen zijn die gedragingen die een redelijke dekking geven van het instructiemodel voor expliciete én impliciete vaardigheden. Gezien het beperkt aantal observaties is geen onderscheid gemaakt tussen de toepassing van de twee afzonderlijke modellen, er is één schaal geconstrueerd. Naast pragmatische redenen is hiertoe ook besloten vanwege de mogelijkheid van vloeiende overgangen tussen beide in-

structiemodellen. In plaats van een tweedeling kunnen beide instructiemodellen gepositioneerd worden als liggend op de uiteinden van een continuüm (Rosenshine & Meister, 1992).

Voorafgaande aan de observaties met de DIS hebben drie observatoren een training van circa 35 uur gevolgd. Bij de training is gebruik gemaakt van video-opnames en zijn op drie niet bij het onderzoek betrokken basisscholen 15 proefobservaties verricht. Vervolgens is aan de hand van vijf lessen in reële klassesituaties de tussen-observator-betrouwbaarheid geschat met behulp van variantie-analyse (Winer, 1971). Voor de twintig afzonderlijke items van de DIS varieert deze betrouwbaarheid van 0.72 tot 1.00 (mediaan 0.96) met als uitzondering het item: 'verwijzing naar nieuwe toepassingssituaties van het geleerde' (0.52).

Op grond van een lesindeling in twee fasen zijn de 20 items van de DIS verdeeld in twee subschalen: presentatie van de leerstof en verwerking van de leerstof (zowel individueel als onder begeleiding). Maten voor interne consistentie (Cronbachs alpha) zijn berekend voor de gehele schaal en voor de twee afzonderlijke subschalen, zowel voor de data verkregen via de getrainde observatoren als voor de data verkregen via de mentoren van de oefenschool. De alpha-coëfficiënten zijn vermeld in Tabel 1 en 2.

De getrainde observatoren hebben tijdens de les die zij observeerden een uitvoerig protocol opgesteld. Aan de hand van dit protocol hebben zij vervolgens de DIS ingevuld.

Nadat de studenten het trainingsprogramma hadden gevolgd, is aan de mentoren van de oefenschool gevraagd of zij ook de DIS wilden invullen. Aan hen is gevraagd of zij terug wilden denken aan de periode vóórdat de studenten het trainingsprogramma volgden en de studenten met behulp van de DIS te beoordelen ('then'-score: retrospectieve voormeting). Vervolgens is hen gevraagd of zij de student wilden beoordelen ná het trainingsprogramma ('post'-score). Het verschil tussen de 'post'-score en de 'then'-score is gebruikt als een index voor de toepassing van de gewenste instructiegedragingen.

4.3 Taakgerichtgedrag van de leerlingen

Om het taakgerichtgedrag van de leerlingen waaraan de studenten op de oefenschool les hebben gegeven te bepalen, is als volgt te werk gegaan. Na 8 minuten de les geobserveerd te hebben stopten de observatoren met het maken van notities en noteerden het aantal leerlingen dat op dat moment taakgericht bezig was. Het verkregen aantal taakgerichte leerlingen werd vervolgens gedeeld door het totale aantal leerlingen dat in de klas aanwezig was. Deze score werd vervolgens omgezet in een percentage om een index te krijgen voor het taakgericht gedrag van de leerlingen. Deze procedure werd toegepast na minuut 8, 16 en 24. Iedere geobserveerde les duurde ongeveer 30 minuten, zodat per les drie schattingen werden verkregen van het taakgerichtgedrag van de leerlingen.

Bij de bepaling van de tussen-observator-betrouwbaarheid is gebruikt gemaakt van dezelfde vijf lessen en van dezelfde procedure als bij de bepaling van de tussen-observator-betrouwbaarheid voor de DIS. De tussen-observator-betrouwbaarheid voor het coderen van het taakgericht gedrag van de leerlingen bedraagt 0.94.

4.4 Gestandaardiseerde lessen

Om de invloed van het trainingsprogramma over effectieve instructie op het onderwijsgedrag van de PABO-studenten onder meer gecontroleerde condities te kunnen vaststellen, zijn twee soorten lesopdrachten samengesteld: taallessen en rekenlessen. Met de constructie van de rekenlessen is beoogd de student uit te nodigen tot het toepassen van het directe instructiemodel voor expliciete vaardigheden. Aan de student is gevraagd optel/aftrekoefeningen en breuken te behandelen. Met de constructie van de taallessen is beoogd de student uit te nodigen tot het toepassen van het directe instructiemodel voor impliciete of hogere cognitieve vaardigheden. In deze lessen is aan de student gevraagd de leerlingen een vaardigheid met betrekking tot begrijpend lezen te onderwijzen. In deze twee soorten lessen is voor de student aangegeven: een korte omschrijving van het doel van de les, de duur van de les (ongeveer 30 minuten), de inhoud van de les en enkele voorbeelden voor de leerlingen om mee te oefenen. Richtlijnen voor het geven van de les zijn niet gegeven.

Deze twee soorten lessen zijn aselect over de studenten van beide onderzoeksgroepen (experimentele en controlegroep) verdeeld.

Iedere student heeft vóór de start van het trainingsprogramma één les op de oefenschool gegeven (voormeting) en één les na afloop van het trainingsprogramma (nameting). Lessen zijn ontwikkeld voor en gegeven aan leerlingen van de groepen 3/4 (onderbouw), 5/6 (middenbouw) en 7/8 (bovenbouw). Voor iedere groep is één taalles en één rekenles ontworpen (in totaal 12 lessen: 6 voor de voormeting en 6 voor de nameting). Studenten die in de voormeting een taalles hebben gegeven, hebben in de nameting ook een taalles gegeven. Dezelfde werkwijze is gevolgd bij de rekenlessen.

4.5 De uitvoering van het trainingsprogramma

In de module 'Effectieve instructie: Leren onderwijzen met behulp van het directe instructiemodel' (Veenman, Leenders, Meyer & Sanders, 1991) worden twee versies van het directe instructiemodel behandeld: een model voor het onderwijzen van expliciete vaardigheden en een model voor het onderwijzen van impliciete of hogere cognitieve vaardigheden. De studenten zijn geïnstrueerd deze twee modellen te gebruiken in hun lessen op de oefenschool. De module bestaat uit drie onderdelen. In deel één komen onderwerpen aan de orde als: de cognitieve ontwikkeling van de leerlingen op de basisschool, de motivatie van de leerlingen om te leren en regels/routines voor goed klassemanagement. Deze onderdelen zijn van belang in de voorbereidingsfase van de les. In deel twee komen de twee directe instructiemodellen aan bod. Dit deel vormt de kern van de module. Het gebruik van de twee instructiemodellen wordt geïllustreerd voor de vakken lezen, rekenen en wereldoriëntatie. Concrete voorbeelden en aanwijzingen worden gegeven voor de toepassing van het directe instructiemodel voor expliciete vaardigheden en het instructiemodel voor impliciete of hogere cognitieve vaardigheden. Tevens wordt een lesvoorbereidingsformulier aangeboden. In deel drie wordt een aantal voorwaarden voor effectieve instructie behandeld. Aan de orde komen onderwerpen als: didactische managementvaardigheden, het voorkomen van probleemsituaties en zelf-evaluatie.

Het trainingsprogramma of de module is gegeven door vier ervaren PABO-docenten. Vóór de start van het programma hebben zij een korte voorscholing ondergaan. Tijdens een workshop van één dag zijn zij geïnstrueerd in de inhoud van de module. Hierbij is getracht het scholingsmodel van Joyce en Showers (1988) te volgen: het aanbieden van de theorie, demonstratie van de instructiemodellen, oefenen met de modellen en feedback over het oefenen. Door gebrek aan menskracht kon de laatste component uit het scholingsmodel, namelijk het geven van coaching van de docenten op de werkplek, niet worden gerealiseerd. Verder is in het kort de handleiding besproken voor het geven van de module, is nader ingegaan op de rol van de docenten bij het onderzoek en is informatie gegeven over de opzet van de evaluatie-studie.

Het aantal lessen dat de PABO-docenten aan de module besteed hebben, varieert van 6 tot 9 lessen (lessen van 90 minuten). De meeste tijd is besteed aan de behandeling van de twee instructiemodellen. Aan de docenten is gevraagd de lessen over de module te geven volgens de beschreven instructiemodellen (modelleren). De vier PABO-docenten hebben de module echter naar eigen interpretatie uitgevoerd. Doordat geen coaching op de werkplek gerealiseerd kon worden, ontstond ruimte voor eigen invullingen. Twee docenten (locatie A en C) hebben gekozen voor het kennisoverdrachtssysteem (doceren) zonder de geselecteerde instructiegedragingen modelmatig voor te doen. Eén docent (locatie B) heeft op grond van een in een later stadium door de PABO gekozen organisatievorm de studenten vooral via zelfstudie de module laten doorwerken. Slechts één docent (locatie C) heeft de uitgangspunten van de twee instructiemodellen modelmatig in de eigen lessen als voorbeeld toegepast.

Gemiddeld hebben de studenten op de oefenschool vijf lessen volgens het directe instructiemodel gegeven. De mentoren zijn door de opleidingen nauwelijks geïnformeerd over de doelen en de inhouden van de module over effectieve instructie. De korte handleiding voor de mentoren, waarin de twee instructiemodellen worden uitgelegd, is nauwelijks gebruikt of geraadpleegd. Slechts een klein aantal mentoren heeft de studenten feedback gegeven over de uitvoering van het directe instructiemodel.

4.6 Dataverzameling- en verwerking

Vóór de aanvang van het trainingsprogramma is elke student uit de observatiestudie tijdens één taal- of rekenles geobserveerd (oktober -

november 1991). Na afloop van het programma is dezelfde student opnieuw tijdens één les geobserveerd (januari - februari 1992). Scores van de reken- en taalles zijn per meetmoment samengenomen om een representatief beeld te krijgen van het lesgeefgedrag van de studenten uit de experimentele en de controlegroep.

De mentoren hebben hun studenten beoordeeld na afloop van het trainingsprogramma (locatie A in december 1991; locaties B en C in februari 1992). In dezelfde periode zijn de evaluatielijsten overhandigd aan de studenten die deel hebben genomen aan het trainingsprogramma. Hierin hebben zij hun oordeel kunnen geven over de inhoud en de opzet van de module. De interviews met de PABO-docenten en de vier experimentele klassen hebben eveneens in deze periode plaats gevonden.

De drie scores voor het taakgericht gedrag van de leerlingen zijn per geobserveerde les gemiddeld om één schatting te verkrijgen voor de klas waaraan de PABO-student heeft lesgegeven. De observatiegegevens van de DIS zijn omgezet in somscores (berekend door de itemscores voor de gehele schaal en voor elke subschaal op te tellen). Bij de toetsing van verschillen tussen voormeting en nameting en tussen experimentele groep en controlegroep is een significantieniveau van 5% aangehouden (eenzijdig). De eenheid van analyse is de PABO-student (of de klas waaraan hij of zij heeft lesgegeven).

Voor een uitgebreide beschrijving van de opzet, het instrumentarium en de dataverzameling wordt verwezen naar Leenders, Meyer en Sanders (1992).

5 Resultaten

Na een vergelijking van de observatiegegevens van de experimentele groep en de controlegroep vóór de start van het trainingsprogramma, worden significante verschillen aangetroffen voor de totaal-score op de DIS, voor de score op de subschaal 'presentatie' en voor het item 'dagelijkse terugblik'. Volgens de observatoren scoren studenten uit de experimentele groep hierop hoger dan studenten uit de controlegroep. Dit verschil kan voornamelijk op rekening worden geschreven van het item 'dagelijkse terugblik'. De studenten uit de experimentele groep beginnen hun lessen vaker met een kort overzicht van hetgeen de vorige les aan de orde was geweest. Met betrekking tot de taakgerichte leertijd van de leerlingen die onderwezen zijn door de PABO-studenten zijn tussen de experimentele groep en de controlegroep tijdens de voormeting geen significante verschillen aangetroffen.

Overwogen is om een covariantieanalyse uit te voeren. Hiervan is afgezien vanwege de mogelijke selectie van de klassen in de experimentele groep door de PABO-docenten (zelfselectie kan correleren met de scores op de voormeting), vanwege het gevonden significante verschil in de DIS-scores tussen de twee groepen op de voormeting en vanwege het relatief kleine aantal studenten.

Om te bepalen of de studenten uit de experimentele groep tijdens de nameting beter gebruik hebben gemaakt van de vaardigheden uit het trainingsprogramma dan tijdens de voormeting, is een *t*-toets voor gepaarde waarnemingen uitgevoerd. Verschillen tussen de experimentele groep en de controlegroep zijn getoetst met behulp van een *t*-toets voor verschilscores (scores op de nameting minus scores op de voormeting).

Tabel 1 bevat een samenvatting van de beschrijvende gegevens voor de DIS volgens de gegevens van de getrainde observatoren, alsmede de resultaten van de uitgevoerde *t*-toetsen.

Uit de gegevens van Tabel 1 blijkt dat de studenten uit de experimentele groep tijdens de nameting significant hoger scoren op de DIS en op de subschalen 'presentatie' en 'verwerking' dan tijdens de voormeting ($p < 0.01$). Dit betekent dat studenten die het opleidingsprogramma hebben gevolgd, tijdens de nameting de instructiegedragingen uit het directe instructiemodel beter toepassen in hun les dan tijdens de voormeting. Voor de studenten uit de controlegroep zijn geen significante verschillen gevonden tussen de voor- en de nameting voor de totale DIS. Wel is er bij de studenten uit de controlegroep een positief significant implementatie-effect gevonden voor de subschaal 'presentatie' ($p < 0.05$). Voor de subschaal 'verwerking' is een negatief significant effect gevonden ($p < 0.01$). Studenten uit de controlegroep maken tijdens de nameting min-

Tabel 1

Overzicht van de gemiddelde scores op de Directe Instructie Schaal (DIS) voor de experimentele en controlegroep volgens het oordeel van de observatoren en de resultaten van de t-toetsen

DIS Subschalen/items	Exp. groep		Controlegroep		Verschilscores		
	Voor	Na	Voor	Na	Exp. groep	Controle groep	t
DIS-totaal (20 items, $\alpha = .86$)	45.7	57.2**	40.7	42.9	11.4	2.2	**
Subschaal Presentatie (11 items, $\alpha = .78$)	27.5	34.0**	23.3	26.8*	6.5	3.1	
Dagelijkse terugblik	2.0	2.8*	1.1	2.2**	.8	1.1	
Activeren voorkennis	2.2	2.7*	1.8	2.4	.5	.6	
Controleren voorkennis	1.8	2.6**	1.4	1.6	.8	.1	*
Aangeven lesdoelstelling	1.6	3.0**	1.6	1.9	1.4	.3	**
Helder taalgebruik	3.9	4.0*	3.9	3.8	.2	-.1	
Het geven van handelingswijzers	2.4	2.6	1.9	1.4	.2	-.5	
Onderwijzen in kleine stapjes	3.5	3.7	3.1	3.4	.2	.3	
Vaardigheid voordoen	2.6	3.2*	1.8	2.7*	.5	.9	
Gebruik concrete voorbeelden	3.3	3.8**	3.0	3.1	.5	.1	
Geven van samenvattingen	1.5	2.3**	1.0	1.4	1.2	.4	**
Gebruik van de DI-structuur	3.0	3.4*	2.9	3.1	.4	.2	
Subschaal Verwerking (geleid en zelfstandig) (9 items, $\alpha = .75$)	22.3	27.0**	21.1	19.52**	4.7	-1.6	**
Controle op begrip/inzicht	2.8	3.0	2.6	2.3	.2	-.3	
Circuleren door de klas	3.3	3.6	3.2	3.0	.3	-.2	
Duidelijke opdrachten	3.8	3.9	3.5	3.5	.1	.1	
Stellen van veel vragen	3.4	3.7	2.9	2.8	.2	-.2	
Wederzijds onderwijzen	1.7	2.1*	1.3	1.2	.4	-.2	*
Leerlingen beginnen meteen aan hun opdrachten	3.0	3.4*	3.3	2.9*	.4	-.5	**
Leerlingen helpen elkaar in kleine groepen	1.4	2.1**	1.4	1.4	.7	-.1	**
Verwijzing naar nieuwe toepassingssituaties	1.3	2.7**	1.5	1.8	1.4	.3	**
Leerlingen verantwoordelijk stellen voor hun werk	1.6	2.5**	1.6	1.5	.9	-.1	**

Opm.: Experimentele groep $n = 27$; Controlegroep $n = 14$. Voor de 4-puntsschaal van elk item geldt: 1 = geen toepassing van de vaardigheid, 4 = duidelijke toepassing van de vaardigheid; * $p < .05$, ** $p < .01$.

der duidelijk gebruik van voor de verwerkingsfase belangrijk geachte instructiegedragingen dan tijdens de voormeting.

Studenten uit de experimentele groep scoren tijdens de nameting op 14 van de 20 items (70%) van de DIS significant hoger dan tijdens de voormeting ($p < 0.05$). Studenten uit de controlegroep scoren op de nameting op slechts 2 items significant hoger, en op 1 item zelfs significant lager ('leerlingen meteen laten beginnen aan hun opdrachten').

Wanneer met behulp van verschilscores de verschillen tussen de experimentele groep en de controlegroep worden onderzocht, blijkt dat studenten uit de experimentele groep significant hoger scoren dan studenten uit de controlegroep op het totaal van de DIS en op de subschaal 'verwerking' ($p < 0.01$). Geen significant implementatie-effect is gevonden voor de subschaal 'presentatie'. Voor drie items van deze subschaal bestaan significante verschillen tussen de experimentele en de controlegroep: controleren of er voldoende voorkennis aanwezig is, duidelijk aangeven van de doelstelling van de les en het geven van een samenvatting van de behandelde stof. Voor de subschaal 'verwerking' zijn bij vijf items significante verschillen aangetroffen tussen de experimentele en de controlegroep: wederzijds of rolwisselend onderwijzen, leerlingen meteen laten beginnen aan hun opdrachten, de leerlingen helpen elkaar in kleine groepen, verwijzen naar nieuwe situaties waarin het geleerde is toe te passen en de leerlingen verantwoordelijk stel-

len voor hun werk.

Met behulp van het programma MANOVA uit SPSS[x] is een aantal univariate variantie-analyses uitgevoerd waarin afzonderlijk de invloed van verschillende controlevariabelen op de verschilscores is onderzocht. Het betreft de (controle)variabelen: PABO (locatie A en B), geslacht van de student, observator (1, 2 of 3), geobserveerde les (rekenen of taal) en de groep van de basisschool waaraan de student heeft lesgegeven (onder-, midden- of bovenbouw). De *F*-waarden zijn berekend met behulp van sequentiële kwadratensommen. Eerst is gekeken naar de totale invloed van de controlevariabele, daarna naar de invloed van de onderzoeksgroep (wel of niet hebben gevolgd van het trainingsprogramma), gecontroleerd voor de controlevariabele, en vervolgens naar de invloed van de interactie tussen de controlevariabele en de onderzoeksgroep. Het blijkt dat de invloed van de onderzoeksgroep op de verschilscores voor de DIS en voor de subschaal 'verwerking' significant blijven als wordt gecontroleerd voor bovengenoemde variabelen. Significante verschillen tussen PABO's, het geslacht van de studenten, de observatoren en de groep waaraan werd lesgegeven zijn niet gevonden. Ook niet tussen reken- en taallessen. Wel is een significant interactie-effect gevonden tussen PABO en onderzoeksgroep voor de subschaal 'presentatie' ($F = 4.56$; $df = 1, 37$; $p < 0.05$). Op locatie A scoren de studenten uit de experimentele groep op deze subschaal tijdens de nameting significant hoger dan de studenten uit de controlegroep. Op locatie B boeken de studenten uit de experimentele groep eveneens een significante vooruitgang, maar de grootste vooruitgang wordt geboekt door de studenten uit de controlegroep. (Op locatie C zijn door de observatoren geen observaties uitgevoerd).

Tabel 2 geeft een overzicht van de scores van de studenten uit de experimentele groep op de DIS en van de resultaten van de *t*-toetsen volgens het oordeel van de mentoren. (Eén item

Tabel 2

Overzicht van de gemiddelde scores op de Directe Instructie Schaal (DIS) voor de experimentele groep volgens het oordeel van de mentoren en de resultaten van de t-toetsen

DIS Subschalen/items	Experimentele groep Voor	Na	*t*
DIS-totaal (19 items, α = .89)	49.7	57.8	**
Subschaal Presentatie (10 items, α = .83)	25.8	30.4	**
Dagelijkse terugblik	2.5	3.1	**
Activeren voorkennis	2.8	3.3	**
Controle voorkennis	2.6	2.9	**
Aangeven lesdoelstelling	2.7	3.0	**
Helder taalgebruik	2.9	3.3	**
Het geven van handelingswijzers	2.4	2.7	**
Onderwijzen in kleine stapjes	2.6	3.2	**
Gebruik concrete voorbeelden	2.8	3.2	**
Geven van samenvattingen	2.3	2.8	**
Gebruik van de DI-structuur	2.4	3.1	**
Subschaal Verwerking (geleid en zelfstandig) (9 items, α = .79)	23.9	27.4	**
Controle op begrip/inzicht	2.9	3.2	**
Circuleren door de klas	3.2	3.6	**
Duidelijke opdrachten	2.8	3.3	**
Stellen van veel vragen	2.6	3.0	**
Wederzijds onderwijzen	1.8	2.1	**
Leerlingen beginnen meteen aan hun opdrachten	3.0	3.2	**
Leerlingen helpen elkaar in kleine groepen	2.7	3.0	**
Verwijzing naar nieuwe toepassingssituaties	2.2	2.7	**
Leerlingen verantwoordelijk stellen voor hun werk	2.7	3.2	**

Opm.: Experimentele groep $n = 41$. Voor de 4-puntsschaal van elke item geldt: 1 = geen toepassing van de vaardigheid, 4 = duidelijke toepassing van de vaardigheid. ** $p < 0.01$.

Tabel 3

Overzicht van de gemiddelde percentages taakgericht gedrag van de leerlingen voor de studenten uit de experimentele en controlegroep volgens het oordeel van de observatoren en de resultaten van de t-toetsen

Groepen onderwezen door PABO-studenten	Exp. groep		Controlegroep		Verschilscores		
	Voor	Na	Voor	Na	Exp.groep	Cont.groep	*t*
Percentage taakgericht gedrag	77	85**	73	68	7.5	-5.4	**

Opm.: Experimentele groep $n = 27$; Controlegroep $n = 14$; ** $p < 0.01$.

- doe de vaardigheid voor - is vervallen ten gevolge van een formuleringsfout). Uit de resultaten van de *t*-toets blijkt dat de studenten na het volgen van het opleidingsprogramma volgens het oordeel van hun mentoren een significante vooruitgang hebben geboekt op de DIS in zijn geheel, op de subschalen 'presentatie' en 'verwerking' en op alle items van de DIS ($p < 0.01$). Voor een kanttekening bij deze gegevens wordt verwezen naar de volgende paragraaf.

Met behulp van het programma ANOVA is een aantal univariate variantie-analyses uitgevoerd om de afzonderlijke invloed te onderzoeken van verschillende controlevariabelen op de verschilscores. Als controlevariabelen zijn gehanteerd: PABO (locatie A, B, of C), de groep van de basisschool waaraan de student heeft lesgegeven (onder-, midden- of bovenbouw), aantal leerlingen in deze groep (minder dan 10, 10 - 20 of meer dan 20), aantal lessen gegeven volgens het directe instructiemodel (geen, 1 of 2, 3 of 4, 5 of 6, meer dan 6) en het gebruik van de mentorenhandleiding bij de bespreking van de lessen (niet, af en toe, regelmatig, elke keer). De variabelen PABO, de groep waaraan is lesgegeven en het aantal leerlingen blijken geen significante invloed te hebben op de verschilscores op de DIS en op de subschalen 'presentatie' en 'verwerking'. De variabele 'aantal lessen gegeven volgens het directe instructiemodel' blijkt daarentegen wel een significante invloed te hebben op de verschilscores voor de subschaal 'presentatie' ($F = 3.30$; $df = 4$; $p < 0.05$). Studenten die méér lessen hebben gegeven volgens het directe instructiemodel boeken een grotere vooruitgang op de subschaal 'presentatie' dan studenten die weinig lessen hebben gegeven volgens het directe instructiemodel. Ook de variabele 'gebruik van de mentorenhandleiding bij de bespreking van de lessen' heeft een significante invloed op de verschilscores van de subschaal 'presentatie' ($F = 3.00$; $df = 3$; $p < 0.05$). Daarnaast heeft deze variabele ook een significant effect op de verschilscores voor de totale DIS ($F = 3.00$; $df = 3$; $p < 0.05$). Het gebruik van de mentorenhandleiding bij de bespreking van de lessen leidt tot meer gerichte feedback en ondersteuning bij het aanleren van de vaardigheden uit het trainingsprogramma, waardoor de studenten deze vaardigheden uiteindelijk op een hoger niveau kunnen toepassen.

Uit Tabel 3 blijkt dat het trainingsprogramma een significant effect heeft op het taakgericht gedrag van de leerlingen van de oefenschool die onderwezen zijn door studenten uit de experimentele groep. Het percentage leerlingen dat taakgericht gedrag vertoont in de experimentele groep is tijdens de nameting significant hoger dan tijdens de voormeting ($p < 0.01$). Tijdens de voormeting bedraagt het percentage leerlingen dat gecodeerd is als taakgericht 77%, in de nameting 85%. Bij leerlingen die les hebben gekregen van studenten uit de controlegroep is daarentegen een afname van taakgericht gedrag geconstateerd: 73% tijdens de voormeting en 68% tijdens de nameting (niet significant). Uit de resultaten in Tabel 3 blijkt tevens dat het verschil in verschilscores tussen de leerlingen uit de experimentele groep en de controlegroep significant is ($p < 0.01$). Dit significante resultaat verandert niet als wordt gecontroleerd voor de variabelen: PABO, geslacht van de student, observator, geobserveerde les en de groep waaraan de student les heeft gegeven. Significante interacties zijn niet gevonden. Het significante verschil in taakgericht gedrag tussen de experimentele groep en controlegroep treedt op voor de scoringsmomenten 2 (na 16 minuten in de les) en 3 (na 24 minuten in de les) en niet voor scoringsmoment 1 (na 8 minuten in de les).

Uit de resultaten van de vragenlijsten en de interviews kan geconcludeerd worden dat het trainingsprogramma door de studenten is gebruikt. De geïnterviewde klassen en de studenten die de vragenlijst hebben geretourneerd, geven aan dat zij het programma zeer bruikbaar

vinden, omdat het veel concrete, specifieke en praktische richtlijnen en suggesties bevat. Vooral de voorbeelden in de tekst zijn waardevol gevonden, omdat zij illustreren hoe de instructiegedragingen uit de twee directe instructiemodellen toepasbaar zijn voor de vakgebieden rekenen, taal/lezen en wereldoriëntatie. Deze positieve beoordeling van het trainingsprogramma kan hebben bijgedragen tot de implementatie ervan.

6 Discussie

De resultaten van het onderzoek duiden aan dat het volgen van het trainingsprogramma 'Effectieve Instructie' een positief effect heeft op het instructiegedrag van PABO-studenten en op het taakgericht gedrag van de leerlingen. Zowel de gegevens van de observatoren als van de mentoren laten zien dat de studenten die het trainingsprogramma hebben gevolgd tijdens hun lessen op de oefenscholen op de gekozen aspecten van het instructiegedrag uit het directe instructiemodel vooruitgang hebben geboekt. De inhoud van het programma is door studenten en docenten positief gewaardeerd.

Hoewel het trainingsprogramma is bedoeld voor eerstejaars studenten van de PABO, is het evaluatie-onderzoek uitgevoerd bij tweedejaars studenten. Hiervoor is gekozen omdat eerstejaars studenten in het eerste semester van het studiejaar nog geen lessen geven op hun stage- of oefenscholen. Op locatie B heeft een beperkt aantal studenten uit het eerste jaar (n=19) het trainingsprogramma toch gevolgd. Uit hun reacties blijkt dat zij het programma iets lager waarderen op het punt van praktische bruikbaarheid dan de tweedejaars studenten op locatie B. Met betrekking tot de bruikbaarheid van het directe instructiemodel voor de taal- en rekenlessen scoren zij ongeveer hetzelfde als de tweedejaars studenten. Over de toepassing van het model bij andere vakken zijn zij minder positief dan de tweedejaars. Ook wordt de praktische bruikbaarheid van het directe instructiemodel voor impliciete vaardigheden door de eerstejaars lager ingeschat.

Uit dit onderzoek blijkt dat de studenten uit de experimentele groep vóór de aanvang van de training hoger scoorden op enkele aspecten van het instructiegedrag zoals vastgesteld met de Directe Instructie Schaal (DIS), met name voor presentatievaardigheden als het geven van een overzicht van de voorafgaande leerstof (dagelijkse terugblik). Omdat hier een significant verschil is aangetroffen ten gunste van de experimentele groep, kan de mogelijkheid van een zekere vorm van zelfselectie niet geheel worden uitgesloten. De PABO-docenten hebben bepaald aan welke klassen zij de module wel of niet zouden geven. Het gegeven dat de studenten uit de controlegroep op locatie B de grootste vooruitgang hebben geboekt ten aanzien van de geselecteerde presentatievaardigheden kan er op wijzen dat de studenten uit deze groep via contacten met de studenten uit de experimentele groep of impliciet via de PABO-docent kennis hebben genomen van de inhoud van de module en deze in hun lessen zijn gaan toepassen. Een zekere mate van 'uitwisseling' kan zijn opgetreden.

Ondanks de korte duur van de cursus (variërend van 9 tot 12 uur), de gebrekkige ondersteuning van de studenten door de mentoren op de oefenscholen en het gebrekkig model staan voor de instructievaardigheden volgens het directe instructiemodel door de PABO-docenten in hun eigen lessen, is het trainingsprogramma 'Effectieve Instructie' succesvol gebleken. Slechts één docent (op locatie C) heeft de in de module beschreven instructiegedragingen in zijn lessen toegepast. Uit de interviews met de studenten blijkt dat zij hierover zeer te spreken waren. Of dit modelleren ook effect heeft gehad op het lesgeefgedrag van de studenten kan niet beoordeeld worden, omdat de observaties op deze locatie door een misverstand geen doorgang vonden.

De inhoud van het programma en de uitvoering ervan kent echter een tweetal beperkingen. Vooreerst is het onderscheid tussen het directe instructiemodel voor expliciete vaardigheden en het directe instructiemodel voor impliciete of hogere cognitieve vaardigheden voor veel PABO-studenten niet duidelijk gebleken. Op één van de opleidingen is alleen het model voor expliciete vaardigheden behandeld. Het leren onderwijzen van hogere cognitieve vaardigheden bleek te moeilijk te zijn. Dit onderdeel van de module over effectieve instructie behoeft verduidelijking. Vormen van ondersteunde instructie (scaffolds) zoals het geven van procedurele steun, het modelleren of voordoen van

de vaardigheid, het illustreren van het maken van keuzen door hardop denken, rolwisselend of wederzijds onderwijzen en het vergroten van de transfermogelijkheden van het geleerde door het aanbieden van gevarieerde toepassingsmogelijkheden dienen concreter te worden uitgewerkt. Verder is gebleken dat de PABO-docenten het voordoen of modelleren van de instructievaardigheden voor hun studenten erg moeilijk vonden. Nodig is een scholing van PABO-docenten waarin de uitgangspunten van het leerling-gezel-meester-systeem centraal staan, dit wil zeggen een scholing waarin PABO-docenten geleerd wordt strategisch om te gaan met de geselecteerde instructiegedragingen en waarin metacognitieve vaardigheden een belangrijke rol spelen. Een dergelijke scholing dient model te staan voor de scholing die de docenten op hun beurt weer aan hun studenten geven. Het systeem van kennisoverdracht dat nu in de voorscholing van PABO-docenten centraal heeft gestaan, heeft de PABO-docenten onvoldoende uitgerust met de (meta)cognitieve kennis en vaardigheden die aan het directe instructiemodel voor impliciete of hogere cognitieve vaardigheden ten grondslag ligt.

De lage respons van de mentoren behoeft enige toelichting. De mentoren zijn pas in een laat stadium in het onderzoek betrokken. Nadat de observaties van de studenten op locatie C niet gerealiseerd konden worden, omdat het trainingsprogramma op het moment van de voormeting reeds gegeven was, werd besloten de mentoren in te schakelen. Op deze wijze zouden alsnog gegevens over het lesgeefgedrag van de studenten op locatie C verzameld kunnen worden. Door ook de mentoren op de locaties A en B om deze gegevens te vragen, konden nieuwe evaluatie-gegevens in het onderzoek worden opgenomen, wat de representativiteit van het onderzoek ten goede zou komen. Door het late tijdstip van inschakeling konden de mentoren onvoldoende op de deelname aan het onderzoek voorbereid worden. Een hiermee samenhangende verklaring voor de lage respons van de mentoren ligt in de toepassing van het gekozen 'then-post' model, een bewerking van het 'pre-then-post' model zoals dit is beschreven door Mezoff (1981). Door de late inschakeling van de mentoren kon een 'pre'-meting niet gerealiseerd worden, het tijdstip van de voormeting was reeds gepasseerd. Na de afronding van de module over effectieve instructie is aan de mentoren gevraagd terug te denken aan het functioneren van de student op het moment van de start van de module en dit functioneren te vergelijken met het functioneren van de student op het moment van afloop van de module. Van de mentoren die de vragenlijst hebben teruggestuurd, hebben 14% de vragenlijst onvolledig teruggestuurd, dit wil zeggen de gegevens betreffende de situatie vóór de module ontbraken. Na drie maanden konden deze mentoren zich niet meer betrouwbaar herinneren hoe de student de geselecteerde instructiegedragingen vóór de aanvang van de module toepaste. Deze onnauwkeurigheid om retrospectief de student te evalueren kan voor veel mentoren een reden zijn geweest om de vragenlijst niet terug te sturen.

Oorspronkelijk heeft het in de bedoeling gelegen de mentoren te benaderen via de studenten die bij hen stage liepen. De vragenlijst waarin de Directe Instructie Schaal was opgenomen, is eerst verzonden naar de opleidingen met het verzoek aan de PABO-docenten om deze lijsten uit de delen aan de studenten uit de experimentele groep en de studenten te vragen de lijst door te geven aan hun mentoren. Op één opleiding is de vragenlijst niet verspreid. Op een andere opleiding hebben de studenten geweigerd om de vragenlijst aan hun mentoren te geven, omdat zij hun mentoren geen extra werk wilden bezorgen. Op de derde opleiding zijn de vragenlijsten alleen uitgedeeld aan de studenten die uitgekozen zijn voor de observaties en niet aan alle studenten die de cursus hebben gevolgd. Hierdoor zijn op deze opleiding een beperkt aantal mentoren bereikt. Door deze communicatiestoornissen is uiteindelijk besloten de vragenlijst per post naar de mentoren te sturen. Tegen de tijd dat de mentoren de vragenlijsten ontvingen, waren sommige studenten echter al overgeplaatst naar andere oefenscholen. Hierdoor hebben veel mentoren het moeilijk gevonden zich te herinneren hoe de student de instructiegedragingen had uitgevoerd.

Niet uitgesloten moet worden geacht dat die mentoren die de vragenlijst wel terug hebben gestuurd een selecte groep vormen. Mogelijk hebben zij duidelijk waarneembare verschillen bij hun studenten geconstateerd. Hierdoor kunnen de resultaten van de mentoren enigszins

geflatteerd zijn. Ook dient gewezen te worden op een mogelijk effect van 'sociale wenselijkheid: door gelijktijdig een score in te vullen voor voormeting én nameting kan de mentor in zijn oordeelsvorming beïnvloed zijn. Dit oordeel kan in positieve zin voor de student zijn uitgevallen.

Ondanks deze tekortkomingen kan gesteld worden dat het trainingsprogramma over effectieve instructie, zoals hier uitgewerkt op basis van onderzoeksbevindingen over effectief onderwijzen, voor de betrokken PABO-studenten en -docenten een succesvolle en waardevolle onderneming is geweest.

Literatuur

ARBO. (1988). *Progressie in Professie: Advies over de ontwikkeling van de PABO*. Zeist: ARBO.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, HJ: Lawrence Erlbaum Associates.

Bickel, W. E., & Bickel, D. D. (1986). Effective schools, classrooms, and instruction: Implications for special education. *Exceptional Children, 52*(6), 489-500.

Brophy, J. E. (1991). Effective schooling for disadvantaged students. In M. S. Knapp & P. M. Shields (Eds.), *Better schooling for the children of poverty: Alternatives to conventional wisdom* (pp. 211-234). Berkeley: McCutchan.

Brophy, J., & Good, T. (1986). Teacher behaviour and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328-375). New York: MacMillan.

Carnine, D. (1989). Direct instruction model: Special Education. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education: Research and studies* (Supplementary Volume One, pp. 251-257). Oxford: Pergamon Press.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction. Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

Creemers, B. P. M. (1991). *Effectieve instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs.

Jong, F. P. C. M. de (1992). *Zelfstandig leren: Regulatie van het leerproces en leren reguleren: een procesbenadering*. Proefschrift Katholieke Universiteit Brabant.

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

Leenders, Y., Meyer, P., & Sanders, M. (1992). *Leren lesgeven met het directe instructiemodel: Een evaluatie-onderzoek naar het trainingsprogramma 'Effectieve instructie'* (Doctoraalscriptie). Nijmegen: Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen.

Mastropieri, M. A., & Scruggs, T. E. (1987). *Effective instruction for special education*. Boston: Little, Brown & Company.

Mezoff, B. (1981). How to get accurate self-reports of training outcomes. *Training and Development Journal, 35*(9), 56-61.

Ministerie van Onderwijs en Wetenschappen (1990). *Perspectief voor de PABO*. Zoetermeer: Ministerie van Onderwijs en Wetenschappen.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117-175.

Parreren, C. F. van (1988). *Ontwikkelend Onderwijs*. Leuven: Acco.

Prawat, R. S. (1989). Promoting access to knowledge, strategy, and disposition in students: A research synthesis. *Review of Educational Research, 59*, 1-41.

Pressley, M., Goodchild, F., Fleet, J., & Zajchowski, R. (1989). The challenges of classroom strategy instruction. *Elementary School Journal, 90*, 301-342.

Resnick, L. B. (1987). Learning in school en out. *Educational Researcher, 16*(9), 13-20.

Roelofs, E., Raemaekers, J., & Veenman, S. (1991). Improving instructional and classroom management skills: Effects of a staff development programme and coaching. *School Effectiveness and School Improvement, 2*, 192-212.

Rosenshine, B., & Edmonds, J. (1990). New sources for improving instruction: The implicit skills study. *Journal of Personnel Evaluation in Education, 4*, 59-73.

Rosenshine, B., & Meister, C. (1992). The use of scaffolds for teaching higher-level cognitive strategies. *Educational Leadership, 49*(7), 26-33.

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

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando: Academic Press.

Simons, P. R. J. (1990). *Transfervermogen*. Inaugurele rede Katholieke Universiteit Nijmegen.

Veenman, S. (1992). Effectieve instructie volgens het directe instructiemodel. *Pedagogische Studiën, 69*(4), 242-269.

Veenman, S. (red.), Leenders, Y., Meyer, P., & Sanders, M. (1991). *Effectieve instructie: leren onderwijzen met behulp van het directe instructiemodel*. Hoevelaken: Christelijk Pedagogisch Studiecentrum.

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

Veenman, S., Lem, P., Roelofs, E., & Nijssen, F. (1992). *Effectieve instructie en doelmatig klassemanagement: Een schoolverbeteringsprogramma voor combinatie- en enkelvoudige klassen*. Amsterdam: Swets & Zeitlinger.

Veenman, S., Lem, P., & Roelofs, E. (1990). Omgaan met combinatieklassen: een onderzoek naar de effecten van een teamgericht nascholingsprogramma. *Pedagogische Studiën, 67*, 45-58.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge: MIT Press.

Walberg, H. J. (1991). Productive teaching and instruction: Assessing the knowledge base. In H. C. Waxman & H. J. Walberg (Eds.), *Effective instruction: Current research* (pp. 33-62). Berkeley: McCutchan.

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

Manuscript aanvaard 1-12-1992

Auteurs

S. Veenman (1942) is als universitair hoofddocent verbonden aan de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

Y. Leenders (1963) is als onderwijskundig medewerkster werkzaam bij het Christelijke Pedagogisch Studiecentrum (CPS) te Hoevelaken.

P. Meyer (1969) studeert onderwijskunde aan de Katholieke Universiteit Nijmegen.

M. Sanders (1966) is als docent onderwijskunde verbonden aan de Hogeschool Katholieke Leergangen te Sittard.

Correspondentie-adres: Vakgroep Onderwijskunde, Katholieke Universiteit Nijmegen, Postbus 9103, 6500 HD Nijmegen; Christelijk Pedagogisch Studiecentrum, Postbus 30, 3870 CA Hoevelaken.

Abstract

Learning to teach according to the direct instruction model: Effects of a preservice teacher preparation program.

S. Veenman, Y. Leenders, P. Meyer & M. Sanders. Pedagogische Studiën, 1993, 70, 2-16.

A quasi-experimental, treatment-control group investigation was designed to test the effects of a preservice training course on effective instruction. Research findings from teacher effects research and cognitive strategy instruction were translated into two direct instructional models: one model for explicit or well-structured skills and one model for implicit skills or higher-level thinking strategies. Following a course on effective instruction from their teacher educators, student teachers implemented standardized lessons to apply selected research-derived teaching behaviours and to increase pupil engagement rates. Based on trained observers' pre- and post-training classroom observations, a significant treatment effect was found for student teachers' teaching behaviours regarding effective instruction and for pupil engagement rates. Ratings from supervising teachers also showed that the student teachers who participated in the course on effective instruction used the recommended instructional skills after completion of the course significantly better than prior to the course.

Onedele motieven voor een edel beroep?*

Een studie naar de motieven om te kiezen voor een hogere pedagogische opleiding

M. Derriks en E. de Kat

Samenvatting

Bij eerstejaars studenten van drie verschillende hogere pedagogische opleidingen zijn de motieven onderzocht. Hiertoe zijn twee vragenlijsten ontworpen. In de ene vragenlijst is de studenten gevraagd naar hun motieven om voor een lerarenopleiding te kiezen, de tweede vragenlijst was gebaseerd op de 'verwachting x waarde-theorie'. Een belangrijk verschil tussen beide vragenlijsten is dat de motievenmeting retrospectief is, terwijl de 'verwachting x waarde-benadering' prospectief is. Verwacht werd dat de 'verwachting x waarde-benadering' minder gevoelig zou zijn voor antwoordtendenties dan de motieven-benadering. Het onderzoek is aan het begin van het studiejaar 1988/89 uitgevoerd onder een aselecte steekproef eerstejaars studenten uit elke opleiding. Op grond van de verzamelde data zijn schalen geconstrueerd. Een vergelijking tussen studenten van de verschillende opleidingen en tussen de beide metingen laat enkele interessante verschillen zien.

Inleiding

In 1989 is door de SCO onderzoek verricht naar de motieven van studenten om te kiezen voor een hogere pedagogische opleiding. De aanleiding voor het onderzoek was de grote zorg bij de hogere pedagogische opleidingen over de 'kwaliteit' van de instromende studenten (Tillema & Veenman, 1985). Met name vanuit de Pedagogische Opleiding Basisonderwijs (PABO) kwamen signalen dat nogal wat studenten de opleiding uit negatieve overwegingen zouden kiezen, en blijk zouden geven van een geringe inzet en studievaardigheden. Dergelijke klachten waren naar de mening van de opleidingen niet nieuw, maar dienden serieuzer genomen te worden, gezien een aantal maatschappelijke ontwikkelingen, zoals de slechte werkgelegenheidsperspectieven voor leraren en de lage aanvangssalarissen.

In het licht van deze ontwikkelingen zou men verwachten dat vooral heel gemotiveerde studenten voor het leraarschap kiezen, terwijl men in de praktijk juist de indruk had dat het daarentegen vooral ging om de zwakkere broeders en zusters die nergens anders terecht kunnen. Voor de lerarenopleidingen was het derhalve van belang te weten wat de instroomkenmerken van de studenten zijn en waarom zij kiezen voor het weinig aantrekkelijke beroep van leraar. De geringe waardering voor het leraarsberoep dateert overigens niet van de laatste jaren, en dat het beroep opzettelijk zo onaantrekkelijk mogelijk gemaakt wordt juist om 'verkeerd' of 'niet' gemotiveerde studenten af te schrikken werd reeds meer dan een eeuw geleden gesuggereerd door de schrijver Marcellus Emants. "Het onderwijs is in Holland een 'voorwerp van aanhoudende zorg der regering' en de geleerde Nederlanders zijn zo begerig de titel van Nederlands onderwijzer te dragen en mede te kunnen werken tot de voortgalopperende ontwikkeling der Nederlandse natie, dat men het onderwijzend personeel voortdurend op harde proeven moet stellen om te onderzoeken of geen onedele drijfveren in 't spel waren toen zij solliciteerden voor een aanstelling. Om dus te voorkomen dat iemand onderwijzer worde in de verwachting van een lui en lekker leven te kunnen leiden is men ertoe moeten overgaan hem een jaargeld toe te kennen, waardoor hij juist voor de hongerdood bewaard wordt, maar waarvan hij zijn vrouw - indien hij zich de

* Het onderzoek is uitgevoerd met subdsidie van de SVO

weelde van een vrouw veroorlooft - nooit een geschenk op haar verjaardag kan geven." Volgens Emants is dit een 'zeer verstandige' strategie, omdat pas wanneer de huizen van de schoolmeesters boven hun hoofden instorten - "eerst dan zullen zij genoeg gelouterd zijn om als leraars der Nederlandse jeugd waardiglijk op te treden, en haar op te leiden 'tot alle christelijke en maatschappelijke deugden'." (Emants, (1877) 1983).

De Nederlandse overheid laat heden ten dage het onderwijzend personeel dan wel niet van de hongerdood omkomen, maar tot voor kort behoorde een ontmoedigingsstrategie gekenmerkt door bezuinigingen (onder andere op de salarissen van beginnende leraren) wel degelijk tot het onderwijsbeleid. Dit maakt de vraag wat studenten momenteel nog doet besluiten om een opleiding tot onderwijsgevende te gaan volgen des te interessanter.

Het onderzoek vond plaats onder studenten van de voltijd opleiding tot leraar in het basisonderwijs (PABO), de voltijd opleiding tot leraar in het voortgezet onderwijs (NLO, ten tijde van het onderzoek nog 'tweevakkig') en de deeltijd opleiding tot leraar in het voortgezet onderwijs (MO).

1 Eerder onderzoek naar motieven

In Nederland is regelmatig onderzoek gedaan naar motieven van studenten om een (bepaalde) wetenschappelijke opleiding te gaan volgen, zowel bij middelbare scholieren, als bij eerstejaars studenten. Dit geldt echter in mindere mate voor de opleidingen in het hoger beroepsonderwijs, en in nog veel mindere mate voor de lerarenopleidingen. Aanwijzingen welke motieven een rol spelen bij de keuze voor een lerarenopleiding zijn echter wel in buitenlandse literatuur te vinden, met name in studies uit de Verenigde Staten. Uit Amerikaans onderzoek wordt duidelijk dat de belangrijkste motieven van studenten om een lerarenopleiding te kiezen in de loop der jaren zijn verschoven. In de jaren vijftig, toen er nog volop werk was, kozen studenten voor het leraarschap vooral vanwege de zekerheid van de baan en de professionele status van het beroep (Fielstra, 1955). In de jaren zestig kozen studenten vooral uit idealistische overwegingen, dat wil zeggen om de samenleving van dienst te zijn of om het onderwijs te veranderen. In de loop van de jaren zeventig en tachtig, als het aantal beschikbare leraarplaatsen in de Verenigde Staten drastisch afneemt, verschuift het accent naar het willen werken met jonge kinderen en het ontplooien van de eigen talenten en mogelijkheden. Tegelijk echter wordt het deel van de studenten dat niet echt een baan in het onderwijs ambieert, groter (Wood, 1978; Allison, 1982; Bergsma & Chu, 1981). Het is interessant na te gaan of de motieven van Nederlandse studenten om een lerarenopleiding te gaan volgen bij deze ontwikkeling aansluiten, juist gezien de overeenkomst in ontwikkelingen wat betreft werkgelegenheidsperspectieven.

2 Theoretisch kader

Wanneer men de motieven van studenten om een bepaalde opleiding te kiezen wil achterhalen is de eenvoudigste methode die de onderzoeker ter beschikking staat, de studenten rechtstreeks te vragen naar het waarom van hun keuze. Deze directe methode is echter niet zonder bezwaren (Drenth, 1960; Härnqvist, 1976; Steltmann, 1980; Van Eck & Houtkoop, 1986). Belangrijke problemen bij het direct vragen naar motieven zijn dat niet iedereen zonder meer bereid is zijn motieven voor een bepaalde keuze aan een onderzoeker prijs te geven. Is iemand daartoe wèl bereid, dan nog krijgt de onderzoeker alleen inzicht in de motieven waarvan die persoon zich bewust is. Omdat de vraag naar iemands motieven vaak achteraf gesteld wordt, is het bovendien niet onwaarschijnlijk dat iemand achteraf eerder zijn reactie op de gemaakte keuze geeft dan zijn beweegredenen daarvoor (Covington, 1983), en dat iemands ervaringen hadat de keuze gemaakt is de antwoorden op de motievenvraag kleuren (Festinger, 1957). Ook zal de persoon aan wie de motievenvraag gesteld wordt vaak geneigd zijn zijn of haar beweegredenen mooier voor te stellen dan ze op het moment van de keuze waren, omdat bij de vraag naar iemands motieven altijd gevraagd wordt naar een verantwoording van gedrag (Hofstee, 1972). Deze bezwaren, verbonden aan het

rechtstreeks vragen naar iemands motieven, betekenen geenszins dat het niet zinvol zou zijn om studenten naar hun motieven te vragen: "perceptions exist among the students, correct or not" (Härnqvist, 1976).

Bovengenoemde bezwaren waren voor ons aanleiding om in dit onderzoek naast een motievenmeting waarbij studenten rechtstreeks naar het waarom van hun keuze gevraagd wordt, een tweede meetmethode uit te proberen, gebaseerd op de 'verwachting x waarde'-theorie. De grondgedachte van het 'verwachting x waarde'-model is als volgt samen te vatten: de keuze van iemand voor een bepaalde handelwijze is afhankelijk van de verwachting dat die handelwijze bepaalde gevolgen zal hebben èn van de waarde die iemand hecht aan die gevolgen (Feather, 1982; Van Geffen, 1977). Respondenten wordt gevraagd vooruit te kijken, zich rekenschap te geven van de gevolgen van hun keuze, en hun waardering uit te spreken over die gevolgen. Hiermee is één belangrijk verschil tussen enerzijds de 'motieven'- en anderzijds de 'verwachting x waarde'-benadering gegeven: de 'verwachting x waarde'-benadering is niet retrospectief, zoals de motieven-benadering, maar prospectief. Respondenten wordt met de 'verwachting x waarde'-benadering niet gevraagd terug te kijken naar het afwegingsproces bij de keuze. Er wordt volstaan met het afzonderlijk expliciteren van verwachtingen en waarderingen ten aanzien van concrete uitkomsten (Nisbett & Ross, 1980). Op grond hiervan verwachten we dat de 'verwachting x waarde'-benadering minder gevoelig is voor de bezwaren die er tegen de motieven-benadering zijn in te brengen en toch een goed inzicht geeft in de beweegredenen van personen.

3 De constructie van de instrumenten

Voor het meten van motieven zijn twee instrumenten ontworpen: een instrument waarin rechtstreeks wordt gevraagd naar de motieven om een hogere pedagogische opleiding te gaan volgen, en een instrument gebaseerd op de 'verwachting x waarde'-theorie. De ontwikkeling van de twee instrumenten vond in twee stappen plaats. Als eerste werd een motievenlijst geconstrueerd en op basis daarvan is de verwachting x waarde-lijst samengesteld.

Voor de ontwikkeling van de motievenvragenlijst zijn interviews met 30 eerstejaars studenten en 12 decanen van alle betrokken hogere pedagogische opleidingen gehouden. De studenten is gevraagd naar de motieven die een rol speelden bij de keuze van de opleiding, het doel dat hen voor ogen stond met het volgen van de opleiding en hun verwachtingen over de opleiding. De in deze gesprekken naar voren gekomen motieven zijn geïnventariseerd, terwijl tegelijkertijd relevante literatuur werd bestudeerd op motieven die bij de keuze voor een lerarenopleiding een rol kunnen spelen. De in de literatuur en de interviews gevonden motieven zijn vervolgens in dertien categorieën ingedeeld die dienden als richtlijn voor de inhoudelijke opzet van de vragenlijst. Een en ander resulteerde in een lijst bestaande uit 62 motieven-items. Gezien de verschillen tussen de drie bij het onderzoek betrokken opleidingen is voor iedere opleiding een afzonderlijke vragenlijst samengesteld met voor een groot deel dezelfde, deels andere motieven. Op een zespuntsschaal (laagste schaalwaarde 0, hoogste schaalwaarde 5) kon worden aangegeven in welke mate het betreffende motief van invloed was op de keuze voor een lerarenopleiding.

Met de motieven-lijst als uitgangspunt is de verwachting x waarde-vragenlijst samengesteld, eveneens bestaande uit 62 items en gedifferentieerd naar de drie opleidingen. In deze vragenlijst gaat het om de verwachte gevolgen van de keuze (het volgen van de opleiding) en de waardering daarvan. Deze vragenlijst is zo geconstrueerd, dat de verwachting x waarde-items inhoudelijk overeenkomen met de motieven-items, zij het dat elk item in de verwachting x waarde-lijst uit twee componenten bestaat: een verwachting- en een waarde-deel. In het waarde-deel kan de student aangeven welke waarde hij hecht aan een bepaald doel dat met het volgen van de opleiding verwezenlijkt kan worden. Verschillende onderzoekers (Vroom, 1964; Van Geffen, 1977) hanteren voor het meten van waarderingen schaalwaarden die lopen van -1 via 0 naar +1, omdat waarderingen van sterk negatief tot sterk positief uiteen kunnen lopen. Naar analogie hiervan hebben wij de waardering gemeten op een 11-punts schaal lopend van -100 via 0 naar +100.

In het verwachting-deel kan aangegeven worden hoe groot de kans (gedefinieerd in procenten) wordt ingeschat dat deze doelen door het volgen van de opleiding daadwerkelijk gerealiseerd worden. Voor het meten van de 'verwachting' is eveneens gebruik gemaakt van een 11-punts schaal, lopend van 0 tot 100% kans. De uiteindelijke motiefscore wordt berekend door deze twee subscores te vermenigvuldigen.

In een proefafname zijn beide vragenlijsten bij een beperkt aantal studenten van elke opleiding afgenomen. Op grond van deze proefafnames zijn enkele items verwijderd, anders geformuleerd of toegevoegd. Het uiteindelijke aantal items in beide lijsten bedroeg 66.

4 De steekproef

Het onderzoek is uitgevoerd onder een representatieve steekproef eerstejaars studenten van de PABO, NLO, en MO. In de MO is het onderzoek beperkt tot de eerstejaars studenten van de tweedegraads opleidingen. Uitgangspunt bij de steekproeftrekking was dat van de drie opleidingen ongeveer evenveel eerstejaars studenten zouden deelnemen. Om zowel vergelijkingen tussen als binnen opleidingen mogelijk te maken moesten per opleiding 250 ingevulde vragenlijsten terug ontvangen worden. Gezien de verwachte nonrespons in de NLO en de MO - als gevolg van problemen bij de verspreiding van de lijsten en het weer inzamelen van de ingevulde lijsten - is besloten een vergrote steekproef te trekken (Van der Linden, 1969). Hoewel dit het nonrespons probleem niet principieel oplost, was het praktisch de enig haalbare oplossing.

Door structurele verschillen tussen de opleidingen (voltijd/deeltijd; algemeen/vakgericht) was het niet mogelijk bij het trekken van de steekproef voor alle vier de opleidingen eenzelfde procedure te volgen. Omdat eerstejaars PABO-studenten een vergelijkbaar onderwijsprogramma volgen is om praktische redenen besloten de afname in dit schooltype klassikaal te regelen. Uit alle 63 vestigingen is een random steekproef van 21 scholen getrokken. Al deze scholen waren bereid om er voor te zorgen dat één willekeurige klas eerstejaars studenten beide lijsten zou invullen. Van de 10 NLO's die er ten tijde van het onderzoek in Nederland waren hebben we er 8, rekening houdend met geografische spreiding, geselecteerd. Alle 8 opleidingen waren bereid om hun medewerking aan het onderzoek te verlenen. De NLO is een vakgerichte opleiding. Om de verschillende vakken evenredig in de steekproef vertegenwoordigd te krijgen zijn per opleiding 60 studenten getrokken uit een door de opleiding ter beschikking gestelde lijst met de namen en vakken van alle eerstejaars studenten. Bij een aantal opleidingen zijn de vragenlijsten naar het huisadres van de studenten gestuurd, een aantal opleidingen was bereid de lijsten zelf onder de studenten te verspreiden. Van de 15 MO opleidingen in Nederland zijn er 11 random getrokken en benaderd. Ook deze 11 opleidingen reageerden positief op het verzoek om medewerking. De verdere procedure verliep analoog aan die in de NLO, alleen ging het hier om maximaal 40 studenten per opleiding.

5 Instrumentafname

Om 'kleuring' van de motieven door ervaringen met het volgen van de opleiding zoveel mogelijk te voorkomen, was het van belang om het tijdsbestek tussen de gemaakte keuze en de motievenmeting zo klein mogelijk te houden (zie theoretisch kader). Dit belang woog voor ons zwaarder dan de mogelijke effecten van de afname van de motieven-vragenlijst op de resultaten van de verwachting x waarde-vragenlijst. Daarom is er voor gekozen direct aan het begin van het studiejaar de motieven-vragenlijst als eerste bij alle studenten af te nemen. Dezelfde studenten kregen zo'n vier weken later de verwachting x waarde-vragenlijst voorgelegd. Door dit tijdsinterval is de verwachting x waarde-lijst door minder studenten ingevuld dan de motieven-lijst. De motieven-lijst is ingevuld door 421 PABO studenten, 313 NLO-studenten en 220 MO-studenten. De verwachting x waarde-lijst is ingevuld door 367 PABO-studenten, 289 NLO-studenten en 159 MO-studenten.

6 Schaalconstructie

Om de verschillende items tot een aantal homogene groepen terug te brengen is per opleiding een exploratieve, hiërarchische clusteranalyse uitgevoerd op de motieven- en verwachting x waarde-variabelen. Er is gekozen voor een hiërarchische clusteranalyse, omdat bij deze methode de classificatie van data stapsgewijs plaatsvindt waardoor zich geen interpretatieproblemen voordoen (ieder item kan slechts in één cluster ingedeeld worden). Bovendien ging het ons niet om het opsporen en/of toetsen van een onderliggend theoretisch model. Bij de clusteranalyse is gebruik gemaakt van de 'average linkage between groups', hetgeen wil zeggen dat de uiteindelijk gevonden clusterstructuur gebaseerd is op informatie over de afstand van alle mogelijke paren items die op basis van de afzonderlijke items gevormd kunnen worden. Op deze wijze zijn negen schalen geconstrueerd met als uitgangspunt dat de motieven- en verwachting- x waarden-schalen dezelfde items zouden omvatten. De meeste van deze schalen komen in alle opleidingen voor, enkele zijn opleidingsspecifiek. Hier volgt een korte inhoudelijke omschrijving van de negen schalen:

Leraar: motieven die betrekking hebben op het lesgeven, een baan in het onderwijs willen.

Toekomst werk: motieven die betrekking hebben op het vergroten van de kansen op de arbeidsmarkt.

Toekomst studie: motieven die betrekking hebben op het volgen van een hogere pedagogische opleiding als opstap naar een andere opleiding.

Eigen ontwikkeling: motieven die betrekking hebben op het vergaren van kennis en vaardigheden om zichzelf verder te ontwikkelen.

Werkomstandigheden: motieven die betrekking hebben op de werkomstandigheden en het aanzien van het beroep van leraar.

Kinderen: motieven die betrekking hebben op het leuk vinden om met kinderen om te gaan.

Opleiding: motieven die betrekking hebben op het karakter en de inhoud van de hogere pedagogische opleidingen.

Vak: motieven die betrekking hebben op kennisvergroting over een bepaald vakgebied dat de kiezer interessant vindt of waar hij (zij) goed in is.

Professionalisering: motieven die betrekking hebben op het vergroten en verder uitdiepen van reeds aanwezige vakkennis om als onderwijzer of leraar beter te functioneren.

In Tabel 1 zijn de alpha coëfficiënten opgenomen van de motieven- en de verwachting x waarde-schalen zoals gevonden in de drie opleidingen. De interne consistentie van de schalen is over het algemeen bevredigend, met uitzondering van de schaal 'opleiding'. De alpha coëfficienten voor deze schaal liggen tussen de .53 en de .62, hetgeen ook voor uitspraken over groepen leerlingen te laag is. Om inhoudelijke redenen is deze schaal echter toch in de verdere analyses meegenomen.

7 Lerarenopleidingen en motieven

De twee door ons uitgevoerde metingen geven in grote lijnen eenzelfde beeld wat betreft de volgorde van belangrijkheid van de verschillende motieven in de drie opleidingen (Tabel 2). In beide metingen zien we dezelfde schalen op de eerste en tweede plaats staan. De PABO-studenten scoren, zowel bij de motieven-als bij de verwachting x waardemeting, gemiddeld het hoogst op de schalen leraar ('een baan in het onderwijs willen') en kinderen ('het leuk vinden om met kinderen om te gaan'). NLO- en MO-studenten scoren in beide metingen gemiddeld het hoogst op de schalen vak ('kennisvergroting op een bepaald vakgebied dat de kiezer interessant vindt of waar hij/zij goed in is') en eigen ontwikkeling ('vergaren van kennis en vaardigheden om zichzelf verder te ontwikkelen').

Er zijn dus duidelijke verschillen in motieven tussen studenten die een opleiding tot leraar in het basisonderwijs gaan volgen en studenten die een opleiding tot leraar in het voortgezet onderwijs gaan volgen. Leraar worden is voor NLO- en MO-studenten bij aanvang van de studie nauwelijks een motief van betekenis om de opleiding te gaan volgen. Bij de PABO- studenten ligt dit duidelijk anders, zij kiezen juist voor de opleiding omdat zij een baan in het onderwijs ambiëren en aantrekkelijk vinden. Van negatieve overwegingen is

Tabel 1

Interne consistentie van schalen per opleiding

		PABO		NLO		MO	
		m	w x v	m	w x v	m	w x v
schaal	*n=*	421	367	313	289	220	159
leraar							
n items		3	3	4	4	4	4
alpha		.88	.90	.96	.96	.88	.89
toekomst werk							
n items		4	4	4	4	6	6
alpha		.60	.50	.68	.67	.72	.72
toekomst studie							
n items		4	4	4	4	4	4
alpha		.79	.73	.65	.68	.70	.77
eigen ontwikkeling							
n items		7	7	7	7	7	7
alpha		.80	.82	.75	.83	.76	.86
werkomstandigheden							
n items		8	8	8	8	8	8
alpha		.76	.78	.85	.89	.81	.87
kinderen							
n items		6	6	6	6	6	6
alpha		.83	.86	.94	.94	.90	.91
opleiding							
n items		6	6	6	6	7	7
alpha		.57	.55	.62	.62	.53	.61
vak							
n items				4	4	4	4
alpha				.61	.81	.65	.69
professionalisering							
n items						4	4
alpha						.74	.78

m = motief; v x w = verwachting x waarde

Tabel 2

Gemiddelde scores op de motieven- en verwachting x waarde-vragenlijst per opleiding

	PABO		NLO		MO	
	m	vxw	m	vxw	m	vxw
leraar	**3.9**	**3.2**	2.0	1.3	1.8	1.3
toekomst werk	2.1	1.9	3.0	2.4	2.3	1.7
toekomst studie	1.0	1.1	1.7	1.9	1.0	1.3
eigen ontwikkeling	2.4	2.9	**3.1**	**3.1**	**3.4**	**3.4**
werkomstandigheden	1.6	2.0	1.2	1.3	0.9	1.1
kinderen	**3.9**	**3.8**	1.8	1.6	1.7	1.6
opleiding	2.3	2.3	2.1	2.0	2.1	2.1
vak			**3.9**	**3.2**	**3.2**	**2.9**
professionalisering					1.6	1.5

* Om de verwachting x waarde-scores vergelijkbaar te maken met de motiefscores is door middel van een lineaire transformatie de maximale schaalscore teruggebracht tot 5.

bij de PABO-studenten dus geen sprake. De opleiding als parkeerstudie gebruiken scoort gemiddeld het laagst. In de NLO en MO daarentegen lijken de eerstejaars studenten over het algemeen niet het plan te hebben leraar te worden. In deze opleidingen lijkt er, afgaande op de bevindingen uit de veldoriëntatie van het onderzoek, meer sprake van een uitgestelde

keuze. Studenten willen geen leraar worden, maar weten ook nog niet wat zij wèl willen. Ook de resultaten van ander onderzoek wijzen in deze richting. In een enquête onder een willekeurige selectie studenten aan de VU (cursusjaar 1980/81) gaf bijna de helft aan indertijd gekozen te hebben uit interesse voor het vak. Bij aanvang van de studie gingen de gedachten nog niet uit naar een later beroep (Prick, 1983). Een onderzoek onder eerstejaars MO-studenten wijst uit dat 25% de studie louter uit interesse is gaan volgen en dat daarenboven bij nog eens bijna de helft de algemene belangstelling zwaarder woog als motief dan de beroepsbelangstelling (MO-overlegraad, 1986; Verheul, Vedder & Bannink, 1989). Deze bevindingen zijn vergelijkbaar met de resultaten van het eerder genoemde Amerikaanse onderzoek uit het begin van de jaren tachtig.

Vergelijking van de rangorde van de overige schalen in beide metingen geeft in de drie opleidingen slechts minimale verschillen te zien, met uitzondering van de schaal 'leraar' in de NLO en de MO. In de motievenmeting komt deze schaal in de beide opleidingen op de vijfde plaats, maar in de verwachting x waardemeting duikelt deze schaal naar een van de laagste plaatsen in de rangorde. Dit betekent dat leraar worden - bij studenten die aan een voltijd of deeltijd lerarenopleiding beginnen - maar heel laag gewaardeerd wordt en dat zij de kans dat zij ook echt leraar zullen worden, uitermate gering achten.

Bij vergelijking van de absolute schaalscores van beide metingen in de drie opleidingen valt op dat de meeste schaalscores in de verwachting x waardemeting lager of gelijk zijn aan de gemiddelde schaalscores in de motievenmeting. In alle drie de opleidingen zijn echter de scores op de schalen 'werkomstandigheden' en 'toekomst studie' in de verwachting x waardemeting hoger dan de schaalscores in de motievenmeting. De schaal 'werkomstandigheden' bevat items die betrekking hebben op de werktijden, het salaris en het aanzien van het leraarsberoep. De schaal 'toekomst studie' bevat items die betrekking hebben op het volgen van een hogere pedagogische opleiding als opstap naar een andere opleiding. Van de items uit beide schalen kan gezegd worden dat ze (tot voor kort in ieder geval) minder aanvaardbaar zijn voor de sociale omgeving als motief om voor een lerarenopleiding te kiezen. Dat juist bij de drie schalen 'werkomstandigheden', 'toekomst studie' en 'leraar' de indirecte meting afwijkt van de directe zou erop kunnen wijzen dat de resultaten van de verwachting x waarde-meting inderdaad minder gevoelig zijn voor een sociaal wenselijke antwoordtendentie.

8 Studievoortgang en motieven

Bij de instellingen zijn de studievoortgang-gegevens verzameld van alle studenten die de twee vragenlijsten hadden ingevuld. Voor de analyse is een driedeling gemaakt naar studenten die het eerste jaar op tijd hebben afgerond (groep 1), studenten die vertraging hebben opgelopen (groep 2) en studenten die zijn uitgevallen (groep 3). In Tabel 3 staat per opleiding een overzicht van de studentenaantallen per groep.

Tabel 3
Studievorderingen per opleiding

	PABO *n*=409	NLO *n*=253	MO *n*=166
groep 1	279 (68%)	94 (37%)	115 (69%)
groep 2	54 (13%)	91 (36%)	7 (4%)
groep 3	76 (19%)	68 (27%)	44 (27%)

groep 1: eerste jaar zonder vertraging voltooid
groep 2: vertraging opgelopen in het eerste jaar
groep 3: studie tijdens of aan het eind van het eerste jaar gestaakt

Met behulp van het SPSS programma 'one-way' is per opleiding geanalyseerd of de drie groepen studenten afwijken op de verschillende schalen van de beide vragenlijsten. Met betrekking tot de motievenmeting zijn alleen bij de PABO verschillen gevonden, en wel bij de twee belangrijkste schalen (zie Tabel 4). De verschillen wijzen in dezelfde richting: degenen die het eerste jaar tijdig hebben afgerond scoren significant hoger (4.0) zowel op de schaal leraar als op de schaal kinderen dan degenen die de studie gestaakt hebben (resp. 3.4 en 3.6). Studenten die de studie al tijdens het eerste jaar afbreken, of er aan het eind van het jaar mee stoppen vinden bij aanvang van de opleiding onderwijzer worden en omgaan met kinderen al minder belangrijk dan studenten

Tabel 4
Gemiddelde motiefscores onvertraagden en studiestakers

		gem.	sd	F-ratio	sign.
PABO					
leraar	groep 1 (*n*=279)	4.0	1.1	6.05	.003
	3 (*n*= 76)	3.4	1.6		
kinderen	groep 1	4.0	0.9	5.62	.004
	3	3.6	1.3		

groep 1: eerste jaar zonder vertraging voltooid
groep 3: studie tijdens of aan het eind van het eerste jaar gestaakt

die uiteindelijk met de opleiding doorgaan.

Bij de andere twee opleidingen hebben we met betrekking tot de motievenmeting geen verschillen gevonden tussen blijvers en studiestakers.

Met betrekking tot de verwachting x waardemeting vinden we zowel bij PABO- als NLO-studenten verschillen tussen blijvers en studiestakers. Bij de PABO gaat het om drie schalen, bij de NLO om één schaal (zie Tabel 5). In de PABO scoren zowel de studenten die volgens programma het studiejaar doorlopen hebben als de studenten die in het eerste jaar vertraging hebben opgelopen, hoger dan de studenten die met de studie zijn gestopt op de twee belangrijkste schalen: 'kinderen' en 'leraar'. Een vergelijkbaar effect vonden we bij de analyse van de motiefscores. Bij de verwachting x waardemeting vinden we ook een verschil op de meest onbelangrijke schaal in de PABO. Op de schaal 'toekomst studie' scoren degenen die met de studie zijn gestopt hoger dan degenen die geheel volgens de planning het eerste jaar hebben doorlopen. Een verklaring voor dit verschil moet naar onze mening gezocht worden in de mogelijkheid een afgeronde propedeuse van de PABO als opstap te gebruiken naar een universiaire opleiding. Een aantal studenten dat aan het eind van het eerste studiejaar de opleiding verlaat, was dit bij de keuze van de opleiding reeds van plan. Dat dit goed verklaarbare effect wel gevonden wordt met de verwachting x waardemeting en niet met de motievenmeting geeft steun aan de veronderstelling dat deze indirecte methode minder gevoelig is voor de eerder, in het theoretisch kader genoemde, vertekening en daardoor de fijne nuances beter zichtbaar maakt.

Ook in de NLO vinden we een verschil tussen blijvers (zowel vertraagden als onvertraagden) en studiestakers dat in de analyse van de directe motievenmeting niet naar voren kwam. Dit verschil betreft een schaal die niet heel belangrijk, maar ook niet geheel onbelangrijk is, nl. de schaal 'opleiding'. De studenten die de opleiding blijven volgen (groep 1 en 2) scoren hoger op deze schaal dan de studenten die de opleiding verlaten (groep 3). Gezien de lage in-

Tabel 5
Gemiddelde verwachting x waardescores onvertraagden, vertraagden en studiestakers

		gem.	sd	F-ratio	sign.
PABO					
leraar	groep 1 (*n*=239)	3.3	1.1	12.65	.000
	2 (*n*= 47)	3.3	1.3		
	3 (*n*= 58)	2.4	1.6		
kinderen	groep 1	3.8	0.8	9.93	.000
	2	3.9	0.8		
	3	3.3	1.3		
toekomst st.	groep 1	1.1	0.3	3.04	.049
	3	1.4	1.3		
NLO					
opleiding	groep 1 (*n*= 83)	2.1	0.8	4.58	.011
	2 (*n*= 74)	2.2	0.8		
	3 (*n*= 56)	1.4	0.9		

groep 1: eerste jaar zonder vertraging voltooid
groep 2: vertraging opgelopen in het eerste jaar
groep 3: studie tijdens of aan het eind van het eerste jaar gestaakt

terne consistentie van deze schaal moeten we erg voorzichtig zijn met de interpretatie van dit verschil. Deze schaal heeft betrekking op kenmerken van de NLO-opleiding, zowel met betrekking tot de inhoud als de organisatie ervan. Studenten voor wie deze motieven een rol bij de keuze gespeeld hebben, hebben blijkbaar wat dat betreft gevonden wat ze zochten.

9 Een nadere vergelijking van de beide metingen

De directe motievenmeting en de verwachting x waardemeting vertonen veel overeenkomsten, maar ook enkele interessante verschillen. Beide metingen geven nagenoeg hetzelfde beeld van het belang van bepaalde redenen om te kiezen voor een hogere pedagogische opleiding, en van de verschillen in motiefsterkte tussen onvertraagde studenten en studiestakers, zij het dat bij de indirecte meting iets meer nuances naar voren komen. Bovendien zijn er aanwijzingen dat de indirecte meting minder gevoelig is voor sociaal wenselijke antwoorden. Deze bevindingen geven in zijn algemeenheid geen aanleiding te concluderen dat de ene meting beter is dan de andere. Welke meting het meest geëigend is, is afhankelijk van de gewenste informatie. Wil men snel op relatief eenvoudige wijze informatie over beweegreden, dan is een motievenmeting voor de hand liggend. Voor een genuanceerder beeld, verdient naar onze mening de verwachting x waarde-benadering de voorkeur.

10 Tot slot

In de inleiding van dit artikel is de 'kwaliteitsproblematiek' van de instroom in hogere pedagogische opleidingen (met name de PABO en de NLO) niet alleen omschreven in termen van motieven van studenten, maar ook in termen van vooropleiding en werkhouding. Dit onderzoek richtte zich op de motieven van instromende studenten en belicht daarmee slechts één, zij het niet onbelangrijk aspect van de 'kwaliteitsproblematiek' van de lerarenopleidingen. Op grond van de resultaten mogen we concluderen dat de pedagogische academies basisonderwijs zich, wat betreft de motivatie van de studenten die zich bij deze instellingen aanmelden, geen zorgen hoeven te maken. Bij de NLO vinden we echter een duidelijke discrepantie tussen de motieven van studenten om de opleiding te gaan volgen en datgene waarvoor de studie opleidt. Net als studenten die een MO-opleiding gaan volgen, zijn algemene vorming, vakinhoudelijke redenen en de gerichtheid op functies buiten het onderwijs belangrijker bij de keuze om de opleiding te gaan volgen dan het leraarschap. Dit levert echter bij de MO een minder grote frictie op dan bij de NLO, omdat MO-studenten na twee jaar de mogelijkheid hebben een algemene studievariant te kiezen. Een antwoord op de vraag of de hogere pedagogische opleidingen nu vooral de zwakkere broeders en zusters binnen krijgen is hiermee niet gegeven. Wel kunnen we vaststellen dat het merendeel van de PABO- en NLO-studenten uit ons onderzoek van het havo afkomstig is (voor beide opleidingen 67%). Het percentage vwo-leerlingen in de PABO is 15%, en in de NLO slechts 9%. Bij de MO-studenten ligt dit anders. Hier is niet meer dan 25% van de instromende studenten afkomstig van het havo, terwijl 20% het vwo heeft doorlopen en 30% in het bezit is van een hbo-diploma. Deze cijfers geven aan dat het percentage vwo-leerlingen dat in de PABO en de NLO instroomt, klein is. Bij vergelijking van de instroom in hogere pedagogische opleidingen met de instroom in andere sectoren van het hoger beroepsonderwijs voor het cursusjaar 1985/86 blijkt bovendien dat de lerarenopleidingen het kleinste percentage vwo-leerlingen binnen krijgen en het grootste percentage studenten met een havo-vooropleiding. Dat de aansluitingsproblemen bij instroom in het HBO voor havo-leerlingen groter zijn dan voor vwo-leerlingen ligt voor de hand. Hun vooropleiding is immers een jaar korter. Wel zijn er aanwijzingen dat, na een aantal jaren ontmoedigingsbeleid van de overheid, vooral de relatief zwakkere havisten naar de lerarenopleidingen doorstromen. Harde uitspraken hieromtrent kunnen op basis van dit motieven-onderzoek niet gedaan worden.

Vakbekwaamheid en inzet van onderwijzend personeel zijn van doorslaggevend belang voor de kwaliteit van het onderwijs en het zou dan ook een slechte zaak zijn als de discrepantie tussen het belang dat algemeen aan het

leraarsberoep gehecht wordt en de maatschappelijke waardering ervoor, geschikte en gemotiveerde studenten afschrikt. Het aantrekkelijker maken van het leraarsberoep kan een bijdrage leveren aan het aantrekken van kwalitatief goede studenten die er bewust voor kiezen om leraar te worden.

Literatuur

Allison, D. J. (1982). *Current attractions of teaching as perceived by preservice students.* Reports-Research/technical (143). Ontario.

Bergsma, H. M., & Chu, L. (1981). *What motivates introductory and senior education students to become teachers?* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Covington, M. V. (1983). Motivated cognitions. In S. G. Paris, G. M. Olson & H. W. Stevenson, *Learning and motivation in the classroom* (pp. 137-164). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Drenth, P. J. D. (1960). *Een onderzoek naar de motieven bij het kiezen van een beroep.* Dissertatie. Amsterdam: Van Soest.

Eck, E. van, & Houtkoop, W. (1986). *Kiezen en begeleiden in het dag/avondonderwijs.* Amsterdam: SCO.

Emants, M. (1983). *Op reis door Zweden.* Amsterdam: Arbeiderspers.

Feather, N. T. (1982). Expectancy-value Approaches: present status and future directions. In N. T. Feather (Ed.), *Expectations and actions* (pp. 395-421). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Festinger, L. A. (1957). *Theory of cognitive dissonance.* Evanstone, IL: Row, Peterson.

Fielstra, C. (1955). An analysis of factors influencing the decision to become a teacher. *Journal of Educational Research, 68,* 659-668.

Geffen, L. M. H. J. van (1977). *De keuze van werk.* Proefschrift. Culemborg: Schoolpers b.v.

Härnqvist, K. (1976). *Individual demand for education: background analytical report.* Parijs: OECD.

Hofstee, W. K. B. (1972). Motieven in de discussie. In P. J. Roscam-Abbing (red.), *Wat drijft of bezielt ons* (pp. 32-47). Meppel: Boom.

Linden, L. A. van der (1969). *Het non-respons verschijnsel bij enquêtes.* Rotterdam: Universitaire Pers Rotterdam.

MO-Overlegraad (1986). *MO-Profiel.* Verslag van een enquête onder 3200 eerstejaars studenten-MO. Utrecht: MO-Overlegraad.

Nisbett, R. E., & Ross, L. (1980). *Human Inferences: Strategies and shortcomings of social judgement.* Englewood Cliffs, NJ: Appleton-Century-Crofts.

Prick, L. G. M. (1983). *Het beroep van leraar.* Amsterdam: VU Boekhandel/Uitgeverij.

Steltmann, K. (1980). Motive für die Wahl des Lehrerberuf. *Zeitschrift für Pädagogik, 26,* 581-586.

Tillema, H. H., & Veenman, S. A. M. (1985). *Probleemverkenning opleiding onderwijsgevenden.* Groningen/Nijmegen/Den Haag: RION/ISO/ITS/SVO.

Verheul, I., Vedder, J., & Bannink, P. (1989). *Instroomveranderingen en studiemotivatie bij eerstejaars studenten van de MO-lerarenvariant.* Utrecht: ISOR/Onderwijskunde/PDI.

Vroom, V. H. (1964). *Work and motivation.* New York: Wiley.

Wood, K. E. (1978). What motivates students to teach? *Journal of Teacher Education, 29*(6), 142-154.

Manuscript aanvaard 13-1-1993

Auteurs

M. Derriks en **E. de Kat** zijn als onderzoeksmedewerker verbonden aan het SCO-Kohnstamm Instituut.

Adres: Grote Bickersstraat 72, 1013 KS Amsterdam.

Abstract

Students' motives to become a teacher

M. Derriks & E. de Kat. Pedagogische Studiën, 1993, 70, 17-27.

To study the motives of students to opt for a career in teaching, an aselect sample of first year students attending three different types of teacher training colleges was drawn. At the beginning of the schoolyear 1988/1989 over a 1000 students filled out two questionnaires. The first questionnaire directly asked for their motives to attend a teacher training college, the second one was based on the expectancy-value theory. An important difference between both methods is that the 'direct' method is retrospective

while the 'expectancy-value' method is prospective. It was expected that the 'expectancy-value' method would be less vulnerable to respons tendencies than the direct method. On the basis of the data collected several scales of motives to attend a teacher training college were constructed. A comparison between the two methods used for measuring motives showed some interesting differences.

Aanvankelijk spellen: de dominantie van de verklankende spelwijze en de geringe effectiviteit van lezen als spellinginstructie-methode

A. M. T. Bosman en M. van Leerdam*

Samenvatting

In dit artikel staan spellinginstructie-methoden en spellingprocessen bij beginnende spellers (basisschoolleerlingen uit Groep 3) centraal. Na een korte training bleken zowel makkelijke als moeilijke woorden overwegend verklankend gespeld te worden. De proefpersonen gingen niet over op een lexicale spellingstrategie. Deze conclusie werd gebaseerd op de resultaten van twee experimenten, waarbij de proefpersonen meer fonologische dan niet-fonologische fouten maakten. Van de hier onderzochte spellinginstructie-methoden was lezen als methode om te leren spellen het minst effectief. De niet-leesinstructie-methoden 'probleem noemen' en 'woord vormen' bleken hetzelfde effect te hebben op de spellingprestaties, maar leidden beide tot betere resultaten dan lezen. 'Probleem noemen' bleek efficiënter te zijn dan 'woord vormen', omdat de proefpersonen in de 'woord vormen' conditie meer trainingstijd nodig hadden dan die in de 'probleem noemen' conditie.

* Met dank aan de leerlingen en leerkrachten van Groep 3 van de scholen Sint Franciscus Xaverius, Don Bosco, De Eenhoorn en Sint Lucia in Haarlem, Sint Antonius en Voorweg in Heemstede, en De Meerpaal, In den Bongerd, De Middelburg, De Rank, Tesselschade en De Vogelboom in Alkmaar.

Tevens dank aan S. E. Gijsberti Hodenpijl, E. A. Jansveld en C. Nederkoorn voor hun bijdrage aan de voorbereiding en assistentie bij de afname van Experiment 1.

De auteurs zijn A. de Groot bijzonder erkentelijk voor haar bijdrage aan de totstandkoming van dit artikel.

1 Inleiding

Spellen blijkt moeilijker te zijn dan lezen. Dat deze asymmetrie (of was het assymetrie) niet beperkt blijft tot leren lezen en spellen wordt geïllustreerd door het feit dat zelfs zeer ervaren taalgebruikers meer woorden correct lezen dan spellen. Ook bij dyslectici blijken spellingproblemen hardnekkiger te zijn dan leesproblemen. Of spellen als vaardigheid inderdaad moeilijker is dan lezen is niet zonder meer vast te stellen, omdat in het onderwijs over het algemeen aan spellen minder aandacht wordt besteed dan aan lezen, en leren spellen van een woord meestal pas gebeurt nadat de leerling blijk heeft gegeven het woord te kunnen lezen.

In dit artikel worden de volgende vragen met betrekking tot leren spellen onderzocht. Ten eerste zal worden nagegaan of het spellingproces van beginnende spellers na een spellingtraining zich ontwikkeld heeft van een verklankende naar een lexicale strategie (uitlezen van het woordbeeld). De veronderstelling is namelijk dat het spellingproces van beginnende spellers zich kenmerkt door het successievelijk omzetten van de fonemen in de daarbij behorende grafemen, terwijl het volwassen spellingproces vooral gekenmerkt wordt door het uitlezen van het woordbeeld uit het mentale lexicon (o.a.: Assink, 1981; Frith, 1980; Sloboda, 1980; Tenney, 1980, maar zie ook Assink, Doorn-van Eijsden & Hees, 1982; Kreiner, 1992).

Ten tweede zal de kwestie van de effectiviteit van verschillende spellinginstructiemethoden onderzocht worden. Hierbij zal met name de diepte van stimulus-verwerking onderwerp van onderzoek zijn. Op de derde plaats zal nagegaan worden of er in een vroeg stadium van het lees- en spellingonderwijs differentiële effecten zijn van verschillende spellinginstructie-methoden.

Dat lezen en spellen verwante processen

zijn zal niemand betwisten, maar in welke mate lezen en spellen elkaars complement zijn is een vraag die onbeantwoord is en overigens ook hier niet afdoende beantwoord zal worden. Op basis van het klassieke 'Dual route model' van visuele woordherkenning zal hieronder uiteengezet worden hoe het spellingproces zich mogelijkerwijs voltrekt. In het 'Dual route model' wordt verondersteld dat het lezen van woorden in een alfabetisch schrift op twee verschillende wijzen kan plaatsvinden. De eerste leeswijze (voornamelijk toegepast door beginnende lezers) is verklanking van de letters in een woord. Dit wordt ook wel spellend, verklankend of indirect lezen genoemd. Deze laatste benaming geeft duidelijk het contrast aan met de tweede manier: namelijk de directe leeswijze (of de lexicale strategie). Hierbij zou geen verklanking meer nodig zijn, maar wordt er een directe match gemaakt tussen het geschreven woord en een mentale representatie. Pas na veelvuldige confrontaties met het woord wordt gebruik gemaakt van de lexicale strategie, ook wel aangeduid met 'het raadplegen van een in het geheugen opgeslagen woordbeeld'. Van ervaren lezers wordt verondersteld dat zij voornamelijk een lexicale leesstrategie volgen. Alleen wanneer zij geconfronteerd worden met laagfrequente woorden gaan zij over op een verklankende leeswijze[1].

Wanneer het uitgangspunt gehanteerd wordt dat lezen en spellen elkaars complement zijn, ligt het voor de hand te veronderstellen dat ook het spellen van een woord op twee manieren kan. Volgens de verklankende of indirecte strategie wordt bij het spellen een woord serieel verklankt, waarbij successievelijk bij elk foneem het juiste grafeem wordt gezocht. Dit lijkt relatief eenduidig, maar dat is het niet, omdat bij een groot aantal fonemen meer dan één grafeem mogelijk is (bijvoorbeeld voor de /ij/-klank bestaan de 'ij' en 'ei' en voor de /g/-klank de 'ch' en 'g'). De grafeem-foneem ambiguïteit is overigens ook aanwezig, maar in mindere mate (het grafeem 'c' kan de /s/- en de /k/-klank representeren). De asymmetrie tussen grafeem-foneem en foneem-grafeem omzetting is eerder opgemerkt door Waters, Bruck en Seidenberg (1985, p. 513). Overeenkomstig de zogenaamde woordbeeld- of lexicale leesstrategie worden bekende woorden gespeld door het raadplegen van een woordbeeld, waarbij de spelling van een woord uitgelezen wordt uit het mentale geheugen.

Uit het voorafgaande kan afgeleid worden dat verklankend spellen door de ambiguïteit in de keuze van grafemen nogal gevoelig is voor fouten. Daarentegen lijkt het aannemelijk dat een strategie waarbij de letters van een woord uitgelezen worden uit het geheugen veel minder vaak tot fouten zal leiden, tenzij de representatie incompleet is of fouten bevat.

Om na te gaan of beginnende spellers na een spellingtraining lexicaal dan wel verklankend spellen zal gekeken worden naar het soort fouten dat door hen gemaakt wordt. Bij het spellen van een woord kunnen de gemaakte fouten verdeeld worden in fouten van fonologische en van niet-fonologische aard. Fonologische spelfouten zijn fouten waarbij het foneem door een verkeerd grafeem wordt weergegeven, maar een dat in zijn klankwaarde gelijk is aan het juiste grafeem, bijvoorbeeld 'haut' in plaats van 'hout'. Bij niet-fonologische fouten is het gekozen grafeem fonologisch niet gelijkwaardig aan het juiste grafeem, bijvoorbeeld 'walt' in plaats van 'valt'. Als bij het spellen van een woord de mentale representatie wordt uitgelezen zal het aantal fonologische fouten even groot moeten zijn als het aantal niet-fonologische. Als daarentegen het aantal fonologische fouten het aantal niet-fonologische blijkt te overtreffen is dit een aanwijzing dat de hier onderzochte steekproef van basisschoolleerlingen voornamelijk een verklankende spelwijze toepast (zie ook: Frith, 1980; Jorm, 1981; Nelson & Warrington, 1974).

De tweede onderzoeksvraag, waarbij naar een mogelijk differentieel effect van lees- en niet-lees-taken (spellinginstructie-methoden) op het aanvankelijk spellen zal worden gekeken, is op de volgende wijze verbonden met de eerste. Er kan bij de hier te onderzoeken spellinginstructie-methoden onderscheid gemaakt worden tussen een oppervlakkige verwerking van woorden en een meer grondige. We veronderstellen dat een grondige verwerking van een woord eerder tot de opbouw van een vollediger representatie leidt dan een oppervlakkige. Tevens wordt aangenomen dat een lexicale spellingstrategie alleen effectief is wanneer de representatie van een woord compleet is. De vraag die zich hier dus opdringt is of

een diepe verwerking van een woord eerder tot een lexicale spelstrategie leidt dan een oppervlakkige. In statistische termen uitgedrukt betekent dit dat er gekeken zal worden of er een interactie bestaat tussen de factor spellinginstructie-methode (zie onder) en de variabele type fout (fonologisch en niet-fonologisch).

Er wordt door sommige onderzoekers vanuit gegaan dat lezen en spellen als twee kanten van één medaille moeten worden gezien (Dumont, 1984, p. 74). Ook Frith (1980) lijkt de mening toegedaan dat lezen en spellen twee complementaire vaardigheden zijn, omdat zij een kind dat een goede lezer is, en tegelijkertijd een zwakke speller aanduidt als iemand met "Unexpected spelling problems" (zie pp. 496-497). Een dergelijke wederkerige relatie wordt ook wel aangeduid met de term "Principle of Associative Symmetry" (Baron, Treiman, Wilf & Kellman, 1980, p. 165), wat wil zeggen dat associaties gevormd in één richting ook bruikbaar zijn in de omgekeerde. Vertaald naar lezen en spellen betekent dit dat iemand die een woord kan lezen dit woord ook kan spellen en omgekeerd.

Uit onderzoek is echter gebleken dat de relatie tussen lezen en spellen *niet* zonder meer voldoet aan the 'Principle of Associative Symmetry' (Baron e.a, 1980; Bosman & De Groot 1991, 1992; Van Doorn-van Eijsden, 1984).

Van Doorn-van Eijsden vergeleek in haar experiment de spellingprestaties van twee groepen leerlingen uit Klas 5 (Groep 7), waarvan één groep de woorden een aantal keren moest lezen en de andere kinderen de woorden dienden over te schrijven. De prestaties van de kinderen in de schrijfgroep waren significant beter dan die van de kinderen in de leesgroep. Ook Bosman en De Groot (1992) vergeleken de spellingprestaties van kinderen die eerder op verschillende manieren een aantal woorden hadden verwerkt. Ook daaruit bleek dat drie niet-lees-methoden betere spellingprestaties tot gevolg hadden dan lezen als spellingtraining. Lezen was als instructie-methode weliswaar effectiever dan helemaal geen instructie, maar het maakte geen verschil in de spelling of kinderen een woord tweemaal of zesmaal hadden gelezen. Dit komt overeen met het eerder gevonden resultaat (Bosman & De Groot, 1991) dat een hoge leesfrequentie (18 aanbiedingen van een woord) niet effectiever was dan een lage (drie aanbiedingen).

Van de drie niet-lees-taken die door Bosman en De Groot (1992) gebruikt werden, bleek 'mondeling spellen' de beste resultaten op te leveren. De instructie-methode 'probleem noemen', waarbij kinderen een ambigu geschreven deel (door onderstreping aangeduid: bijvoorbeeld de 'ei' in paleis) van een woord hardop moesten benoemen, bleek even effectief (de score op de spellingtest was gelijk aan die van 'overschrijven'), maar efficiënter te zijn dan de instructie-methode 'overschrijven', omdat de trainingsprocedure voor 'probleem noemen' significant minder tijd in beslag nam dan die voor 'overschrijven'. Met andere woorden: het was dus niet de duur maar de kwaliteit van de geïnvesteerde tijd die een verklaring vormde voor de betere spellingprestaties.

Het onderzoek van Bosman en De Groot (1992) naar differentiële effecten van spellinginstructie-methoden werd uitgevoerd in de maanden mei en juni, na gemiddeld negen maanden lees- en spelonderricht. Een deelvraag die in het hieronder te rapporteren experiment onderzocht werd was of de niet-leestaken bij een vergelijkbare groep basisschoolleerlingen (Groep 3) ook in een zeer pril stadium effect sorteren. Daartoe zal nagegaan worden of beginnende lezers/spellers in februari al een differentieel effect ondervinden van verschillende spellinginstructie-methoden, waaronder lezen.

Voordat we overgaan tot de beschrijving van de experimenten wordt eerst een beeld geschetst van de wijze waarop de spelling van woorden in Groep 3 geleerd wordt. De plaats van spelling in de leergang "Veilig leren lezen", waarin de hier participerende proefpersonen onderwezen worden, is ondergeschikt aan die van lezen. Volgens de methode moet spellen niet te zeer geïsoleerd worden en behoeven lezen en spellen geen gelijke tred te houden; het herkennen van een geschreven woord is gemakkelijker dan de reproduktie ervan (Caesar, 1980). Er wordt ook gesteld dat door lezen vanzelf een woordbeeld gevormd *kan* worden, maar dat lezen alleen niet altijd voldoende is. De aanbevolen werkwijze voor het leren van de schrijfwijze van een woord in het eerste leerjaar is vooral het inprenten van woordbeelden door middel van het tonen van een woord (zo-

genaamde 'woordstroken'), waarna het kind het woord 'in de lucht' moet schrijven, na moet leggen met behulp van de letterdoos en later in het jaar moet overschrijven.

De eerste woorden die kinderen in deze methode leren lezen en spellen zijn klankzuiver, zodat beheersing van de grafeem-foneem correspondentieregels in principe voldoende is om deze te schrijven. De spelling van de woorden die onderwezen worden in het eerste jaar van het reguliere lees- en spellingonderwijs is nauwelijks of niet beregeld, een uitzondering hierop is de finale d[2].

2 Experiment 1

In Experiment 1 stonden drie vragen centraal: 1) Zijn er aanwijzingen dat het spellingproces van beginnende spellers zich ontwikkelt van een verklankende naar een lexicale strategie? 2) Wat is de effectiviteit van een aantal in diepte van verwerking te onderscheiden spellinginstructie-methoden en leiden de onderzochte spellinginstructie-methoden tot verschillende spellingstrategieën? 3) Zijn de verschillende instructie-methoden onderwijsbaar in een vroeg stadium van het systematische lees- en spellingonderricht?

2.1 Methode

Ter introductie van de experimentele opzet zal eerst een globale beschrijving volgen van de procedure die een proefpersoon doorliep. Aan een kind werd gevraagd om vijf woorden die het kind waarschijnlijk niet kende enkele malen hardop voor te lezen. Daarnaast moest het kind van vijf andere woorden de spelling oefenen. Het oefenen van de spelling kon op vier verschillende manieren plaats vinden (zie opzet en procedure). Elk kind leerde de spelling van de woorden met één van de vier methoden. Tijdens de training werd bijgehouden hoeveel fouten het kind maakte en tevens hoeveel tijd de training in beslag nam. Na de lees- en spellingtraining werd een schriftelijk dictee afgenomen over de tien geoefende woorden. De dicteewoorden werden op twee verschillende manieren beoordeeld. Op de eerste plaats werd nagegaan of een van tevoren bepaald grafeem correct was (target-beoordeling, zie onder) en ten tweede werd het aantal fonologische en niet-fonologische fouten bepaald (type fout beoordeling, zie onder). In bijlage A staat het onderzoeksdesign schematisch weergegeven.

Materiaal

Met behulp van de 'Nieuwe streeflijst woordenschat voor 6-jarigen' van Kohnstamm, Schaerlaekens, De Vries, Akkerhuis en Frooninckx (1981) werden 10 woorden gekozen (gemiddelde lengte 3.9 letters) die niet behoren tot de basisleesstof van Groep 3 van scholen die gebruik maken van de leesmethode "Veilig leren lezen" (Caesar, 1979). Deze woorden, die bij kinderen van 6 jaar als bekend mogen worden verondersteld, vormden het stimulusmateriaal van Experiment 1[3].

De stimuli werden gekozen op grond van een verwachte moeilijkheid in de schrijfwijze. Dit betrof de aanwezigheid van een fonologisch ambigue letter of lettercluster in een woord, zoals bijvoorbeeld de 'ei' in 'reis'. Hiervoor vormt het fonologisch identieke 'ij' een mogelijk alternatief. In Bijlage B staan de gebruikte stimuli[4]. Bij elke stimulus staat tevens aangegeven welke letter of lettercluster (in het vervolg 'target' genoemd) mogelijk moeilijkheden zal opleveren bij de spelling. De targets vormden tevens de basis voor beoordeling van de responsen.

De tien stimuli werden opgedeeld in twee groepen van elk vijf woorden. De woorden 1 tot en met 5 vormden de eerste groep en 6 tot en met 10 de tweede. Door deze indeling werd ervoor gezorgd dat in beide groepen dezelfde targets voorkwamen. In elke groep kwam elk woord drie keer voor, waarbij de volgorde van de woorden zodanig was dat geen enkel woord twee maal achter elkaar verscheen (in elke conditie werden dus 15 woorden geoefend). Om te voorkomen dat één groep stimuli uniek gekoppeld werd aan één conditie werd aan de helft van de proefpersonen de ene groep woorden in de lees- en de andere groep woorden in de niet-lees-conditie aangeboden (zie opzet en procedure). Bij de andere helft van de proefpersonen werd dit omgedraaid. Hiermee werd tevens bewerkstelligd dat een stimulus uitsluitend in één van de twee condities werd gepresenteerd en nooit in beide.

Elk woord werd apart op een kaartje van vijf bij tien centimeter afgedrukt in het lettertype Helvetica. Dit type drukletter is voor de kin-

deren van de onderzochte leeftijdsgroep zeer bekend, omdat het wordt gebruikt in de de eerste elf deeltjes van de leesmethode "Veilig leren lezen".

Opzet en procedure

Een complete experimentele sessie, waarbij elk kind individueel werd getest, duurde gemiddeld 15 minuten. De kinderen kregen in februari een spellingtraining, waarbij ze toegewezen werden aan één van de vier volgende condities (niet-lees-taak): 'overschrijven', 'mondeling spellen', 'probleem noemen', of 'woord vormen'. Naast de training van 15 woorden in een niet-leestaak oefenden de kinderen ook 15 (andere) woorden in de leestaak. De leestaak bestond uit het hardop lezen van de woorden. Wanneer een woord fout opgelezen werd verzocht de proefleider het kind het woord nogmaals te lezen. De helft van de kinderen kreeg de niet-lees-taak in de 'stimulusafwezig'-conditie en de andere helft kreeg deze in de 'aanwezig'-conditie (zie onder).

In de 'overschrijven'-conditie moest een kind de woorden opschrijven in een kladblok, waarbij voor elk woord een aparte bladzijde werd gebruikt. Voordat het woord werd overgeschreven werd aan het kind gevraagd het woord eerst eenmaal hardop voor te lezen. De kinderen in de 'aanwezig'-conditie schreven de woorden over van de kaartjes, terwijl bij de kinderen in de 'afwezig'-conditie de woorden tijdens het opschrijven werden afgedekt.

In de 'mondeling spellen'-conditie werd een kind gevraagd de woorden eerst hardop voor te lezen en deze daarna uit het hoofd ('afwezig'-conditie) hardop te spellen of vanaf het kaartje hardop te spellen ('aanwezig'-conditie). Omdat kinderen uit Groep 3 geneigd zijn om bij het analyseren van woorden de fonemen op te noemen in plaats van de letters, werd hen bij elk dubbelzinnig foneem gevraagd dit te verduidelijken.

In de 'probleem noemen'-conditie diende een kind een woord ook eerst hardop voor te lezen, waarna het een kaartje kreeg aangeboden waarop twee letters of letterclusters stonden waarvan er één de target was. Bijvoorbeeld, bij het woord 'paleis' werd een kaartje aangeboden met de letterclusters 'ij' en 'ei'. Aan het kind werd gevraagd om met een potlood de correcte letter of lettercluster te omcirkelen. De plaats ervan op het kaartje (aan de linker- of rechterkant) werd bepaald door middel van pseudorandomisatie. De kinderen die binnen de 'probleem noemen'-groep aan de conditie 'afwezig' waren toegewezen kregen het kaartje met het woord slechts éénmaal te zien, terwijl de kinderen in de 'aanwezig'-conditie er voortdurend naar mochten blijven kijken.

Ook in de 'woord vormen'-conditie moest een woord eerst worden voorgelezen. Vervolgens werd aan het kind gevraagd om het woord te vormen met behulp van letterblokjes. Aan deze letterblokjes werd, behalve de letters waaruit het woord bestond, een letter of lettercluster toegevoegd die een mogelijk alternatief vormde voor de target van het woord. Voor bijvoorbeeld het woord 'zoen' werden dus niet alleen de letters en lettercluster 'z', 'oe' en 'n' gegeven, maar ook de letter 's'. Net als in de andere condities gold dat de kinderen die binnen de 'woord vormen'-groep aan de conditie 'afwezig' waren toegewezen het kaartje met het woord slechts éénmaal te zien kregen, terwijl de kinderen in de 'aanwezig'-conditie het kaartje voortdurend voor zich hadden liggen.

De ene helft van de kinderen nam eerst deel aan de leestaak alvorens een van de vier niet-leestaken uit te voeren, voor de andere helft was de volgorde omgekeerd. De tijd die een kind nodig had voor het uitvoeren van zowel de lees- als de niet-lees-taak werd met een stopwatch opgemeten.

Direct na de training werd een dictee afgenomen. Voor elk van de 10 woorden was een bijpassende zin geconstrueerd. De proefleider las iedere zin één keer voor en herhaalde daarbij telkens eenmaal het woord dat door het kind moest worden opgeschreven.

Proefpersonen

Bij een steekproef van 170 leerlingen uit Groep 3 van basisscholen die de leesmethode "Veilig leren lezen" van Caesar (1979) gebruiken, werd een week voor de afname van het experiment een vooronderzoek uitgevoerd. Dit vooronderzoek bestond uit de Eén-Minuut-Test (Caesar, 1970) en een door de eerste auteur ontworpen spellingtest[5]. In Bijlage C staat de spellingtest afgedrukt. De testwoorden zijn onderstreept.

Op basis van de twee testscores werd voor elk kind een gemiddelde *Z*-score berekend. Van de 149 kinderen die deelnamen aan het experimentele gedeelte van het onderzoek werden er uiteindelijk 136 opgenomen in de analyses. De gemiddelde scores van de proefpersonen waren 17.07 (*SD* 2.7) op de spellingtest (maximum score is 20) en 30.75 (*SD* 11.4) op de leestest (maximum score is in principe oneindig). De significante correlatie tussen de scores op de lees- en spellingtest bedroeg .34 ($p < .01$).

De opname van proefpersonen in de analyses geschiedde op grond van twee voorwaarden. De eerste voorwaarde was dat er zich in elke conditie een gelijk aantal kinderen ($n = 17$) bevond. De tweede was dat er geen significant verschil mocht bestaan tussen de experimentele condities op de gemiddelde *Z*-score op de lees- en spellingtest ($F(7,128) = .006, p = 1.0$). Verschillen tussen condities kunnen dus niet verklaard worden door verschillen in lees- en spellingvaardigheid.

2.2 Resultaten

De bespreking van de resultaten wordt hier opgesplitst in twee delen. Eerst worden de gegevens van de trainingsfase behandeld. Daarna komen de resultaten van het dictee (de testfase) aan de orde. Op basis van de spelfouten in het dictee kon het effect van de experimentele behandeling worden geëvalueerd. De dicteewoorden werden op twee wijzen beoordeeld. De eerste betrof de aan- of afwezigheid van de zogenaamde target (zie materiaal sectie), waarbij de beoordeling dichotoom was namelijk, 1 (goed) of 0 (fout). Een fout in het niet-target deel van het woord was in deze wijze van beoordelen irrelevant.

Bij de tweede wijze van beoordeling werd de schrijfwijze van het totale woord in beschouwing genomen, waarbij nagegaan werd of gemaakte fouten fonologisch dan wel niet-fonologisch zijn. Spelfouten in woorden werden geclassificeerd als 'fonologisch' wanneer door het toepassen van letter-klank correspondentieregels het woord kon worden uitgesproken zoals het correct gespelde woord. In het fout gespelde woord 'flag', bijvoorbeeld, wordt de 'f' beoordeeld als een fonologische fout, omdat het woord op dezelfde wijze uitgesproken wordt als 'vlag'. Spelfouten als 'friut' ('fruit') of 'veis' ('vies') waarbij fonemen waren weggelaten, vervangen of omgewisseld, werden beoordeeld als 'niet-fonologisch'[6]. Dit betekent dat in tegenstelling tot de target-beoordeling het gemiddeld aantal fouten per woord meer dan 1 kon bedragen (de maximum foutscore is onbepaald). Tevens is de objectieve kans op fonologische fouten kleiner dan op niet-fonologische, er zijn immers gemiddeld minder fonologisch ambigue grafemen dan niet-ambigue.

Trainingsfase

Een 4 (niet-lees-conditie: overschrijven vs. mondeling spellen vs. probleem noemen vs. woord vormen) bij 2 (stimulusaanwezigheid: afwezig vs. aanwezig) bij 2 (taak: lezen vs. niet-lezen) variantie-analyse op de trainingsduur met herhaalde metingen op de variabele taak liet significante hoofd- en interactie-effecten zien. Het hoofdeffect van taak was $F(1,128) = 1193.64, p < .001$. Het uitvoeren van de leestaak nam significant minder tijd in beslag (gemiddeld 1.16 min) dan het uitvoeren van de niet-lees-taak (4.18 min). Ook het hoofdeffect van de niet-lees-conditie was significant ($F(3,128) = 47.23, p < .001$). Uit een Newman-Keuls post hoc analyse op een niveau van minimaal 5% bleek dat de gemiddelde tijdsduur om een instructie uit te voeren voor alle vier de niet-lees-niveaus significant van elkaar verschilde (mondeling spellen = 1.79, probleem noemen = 2.36, overschrijven = 2.78 en woord vormen = 3.76 min). Het hoofdeffect van stimulusaanwezigheid was marginaal significant, $F(1,128) = 3.66, .05 < p < .10$. De spellingtraining in de 'afwezig'-conditie nam meer tijd in beslag dan in de 'aanwezig'-conditie (resp. 2.79 en 2.55 min.). Het significante interactie-effect ($F(3,128) = 82.04, p < .001$) tussen taak en niet-lees-conditie toonde aan dat de tijden nodig om de leestaak te voltooien voor de vier verschillende niveaus van de lees-conditie niet significant van elkaar afweken, maar die voor het voltooien van de vier niet-lees-taken wel ($F(3,128) = 70.59, p < .001$).

Een variantie-analyse op het aantal fouten tijdens de training gaf aan dat er geen statistisch significant verschil bestond tussen de vier niet-lees-condities (spellinginstructie-methoden). Wel werden tijdens de niet-leestaak (.81) meer

fouten gemaakt dan tijdens de leestaak (.54), ($F(1,128) = 6.00, p < .05$).

Testfase

Uit een 4 (niet-lees-conditie: overschrijven vs. mondeling spellen vs. probleem noemen vs. woord vormen) bij 2 (stimulusaanwezigheid: afwezig vs. aanwezig) bij 2 (taak: lezen vs. niet-lezen) variantie-analyse op het aantal targetfouten met herhaalde metingen op de variabele taak bleek geen enkel hoofd- of interactie-effect significant.

Uit een zelfde analyse als hierboven, maar met de factor type spelfout (fonologisch vs. niet-fonologisch) toegevoegd bleek alleen een significant hoofdeffect van type fout $F(1,128) = 78.83, p < .001$. Het gemiddeld aantal fonologische fouten (.25) lag significant hoger dan het aantal niet-fonologische fouten (.09). Gemiddelde target-, fonologische en niet-fonologische fouten zijn weergegeven in de Tabellen 1 en 2 in Bijlage D.

2.3 Conclusie

De resultaten van het eerste experiment zijn *niet* in overeenstemming met die van Van Doorn-van Eijsden (1984) en Bosman en De Groot (1992). Daaruit was gebleken dat woorden geleerd in één van de niet-lees-taken betere spellingresultaten opleverden dan woorden in de leestaak. In Experiment 1 was er echter geen verschil tussen de spellingprestaties in beide condities.

Ook was er, in afwijking van het onderzoek van Bosman en De Groot, geen differentieel effect van de vier verschillende spelling-instructie-methoden. Alle vier methoden (overschrijven, mondeling spellen, probleem noemen en woord vormen) leverden dezelfde spellingresultaten op. De door ons toegepaste spellinginstructie-methoden verschilden wat diepte van verwerking betreft. We veronderstelden dat 'mondeling spellen' en 'woord vormen' een grondiger verwerking van het woord vereisten dan overschrijven en probleem noemen. Het feit dat er geen verschillen tussen condities optraden lijkt aan te geven dat een verschil in diepte van verwerking van de stimulus geen invloed heeft op de spellingprestaties. De vraag of kinderen in een zeer vroeg stadium van het lees- en spellingonderwijs al baat hebben bij de hierboven geschetste spelling-instructie-methoden lijkt dus in eerste instantie met nee beantwoord te moeten worden. Maar een waarschijnlijk niet onbelangrijk verschil tussen Experiment 1 en de studie van Bosman en De Groot zou het uitblijven van het effect ook kunnen verklaren.

De moeilijkheidsgraad van de stimuli uit Experiment 1 was namelijk aanzienlijk geringer dan die van Bosman en De Groot. Hun stimuli waren vergelijkbaar met die van de woorden in Experiment 2 (hieronder). Een alternatieve verklaring voor het uitblijven van het effect is dus dat er sprake geweest kan zijn van een plafond-effect. Deze mogelijke verklaring wordt door twee gegevens ondersteund. Op de eerste plaats was de proportie fout in Experiment 1 (lees-taak: .28, niet-lees-taak: .26) vergeleken met die van Bosman en De Groot (lees-taak: .96, niet-lees-taak: .57) relatief laag. Ten tweede had de variabele stimulusaanwezigheid tijdens de training geen effect op de prestaties op het dictee. De kinderen in de stimulus-afwezige conditie (.25) maakten niet meer fouten dan kinderen in de stimulus-aanwezige conditie (.29).

De trainingsresultaten van Experiment 1 komen gedeeltelijk wel overeen met die van Bosman en De Groot (1992). Uit beide onderzoeken bleek dat de leestaak minder tijd in beslag nam dan de niet-lees-taken. Omdat de variabele stimulus-aanwezigheid in het onderzoek van Bosman en De Groot niet gemanipuleerd werd, en tevens de instructie-methode 'probleem noemen' wat uitvoering betrof niet geheel overeen kwam met die van Experiment 1, zijn de trainingstijden voor de gebruikte instructie-methoden van Experiment 1 echter niet vergelijkbaar met die van Bosman en De Groot.

De niet-lees-taken bleken niet in moeilijkheidsgraad te verschillen (geen significant verschil in het aantal fouten tijdens de training), maar de kinderen vonden de niet-lees-taken wel moeilijker dan de leestaak (meer fouten in de niet-leestaak dan in de leestaak tijdens de training).

Uit het feit dat de kinderen meer fonologische dan niet-fonologische spelfouten maakten in het dictee zou geconcludeerd kunnen worden dat zij voornamelijk verklankend spellen. Het uitblijven van interacties tussen de variabele type fout en niet-lees-conditie, tussen type

fout en taak (lezen vs. niet-lezen) en tussen de variabele stimulusaanwezigheid en type fout is een aanwijzing dat de onderzochte spellinginstructie-methoden niet tot verschillende spellingstrategieën leiden. Omdat het echter niet duidelijk is of de gebruikte stimuli te makkelijk waren zijn deze conclusies op dit moment nog niet gerechtvaardigd.

Het bovenstaande gaf aanleiding om Experiment 1 gedeeltelijk te repliceren. Om na te gaan in hoeverre er sprake was van een plafondeffect werd naast de stimuli uit Experiment 1 een aantal moeilijker te spellen woorden geselecteerd en een controle-groep toegevoegd.

Omdat uit de studie van Bosman en De Groot (1992) duidelijk was gebleken dat 'mondeling spellen' de beste methode om te leren spellen was en 'overschrijven' de minst effectieve, werden deze methoden in Experiment 2 niet meer in het onderzoek betrokken. De instructie-methode 'woord vormen' was niet onderzocht in het onderzoek van Bosman en De Groot en werd daarom gehandhaafd in Experiment 2. Om de vraag te kunnen beantwoorden of een grondige verwerking van een woord tot een kwalitatief andere, vaker lexicale, spellingstrategie leidt, werd 'probleem noemen' opnieuw opgenomen. Het belangrijke verschil tussen 'probleem noemen' en 'woord vormen' betreft de volledigheid van de verwerking. Bij 'probleem noemen' hoeven de kinderen alleen te letten op de target, terwijl bij 'woord vormen' het hele woord gereconstrueerd moet worden. De factor stimulusaanwezigheid werd in Experiment 2 niet gemanipuleerd; er is gekozen voor afwezigheid van de stimulus tijdens de training. Dit komt meer overeen met de uitvoering van het onderzoek van Bosman en De Groot (1992).

De afname van Experiment 2 vond plaats in mei, drie maanden na het eerste. De vraag naar het effect van de hier onderzochte spellinginstructie-methoden in een zeer vroeg stadium kan hier dus niet meer opgelost worden, maar de beantwoording van de twee andere vragen (zie inleiding Experiment 1) is met onderstaand experiment wel mogelijk.

3 Experiment 2

3.1 Methode

De gevolgde procedure is identiek aan die van Experiment 1.

Materiaal

Het stimulusmateriaal bestond uit de tien woorden van Experiment 1 (makkelijke woorden) en tien nieuwe woorden (moeilijke woorden; zie Bijlage B). Deze moeilijke woorden (gemiddelde lengte 6.7 letters) werden op dezelfde wijze geselecteerd als de makkelijke woorden ('Nieuwe streeflijst woordenschat' van Kohnstamm e.a., 1981). De moeilijke woorden onderscheiden zich van de makkelijke woorden door een grotere woordlengte en een groter aantal potentiële moeilijkheden in de schrijfwijze. Op dezelfde manier als in Experiment 1 werden de tien nieuwe woorden opgedeeld in twee groepen van elk vijf waarbij elk woord weer drie keer verscheen. Nu werden de woorden echter, in plaats van ieder woord apart op een kaartje, afgedrukt op een vel papier op A-4-formaat met één woord per regel. De bedoeling van deze verandering was om de training enigszins praktischer te doen verlopen.

Opzet en procedure

Net als in Experiment 1 werden de kinderen individueel getest. Ook nu hadden zij ongeveer 15 minuten nodig om de experimentele sessie af te ronden. De opzet en procedure waren in grote lijnen gelijk aan die van Experiment 1. Opnieuw voerde elk kind zowel een leestaak als een niet-lees-taak uit. In tegenstelling tot Experiment 1 werden in Experiment 2 slechts twee van de vier niet-lees-taken opgenomen, namelijk 'probleem noemen' en 'woord vormen'. De helft van de kinderen kreeg de training met de makkelijke woorden en de andere helft oefende de moeilijke woorden. De niet-lees-taak vond plaats volgens dezelfde procedure als die in de 'stimulus-afwezig'-conditie van Experiment 1. In deze nieuwe opzet moesten alle kinderen dus het woord eenmaal hardop voorlezen waarna gedurende de periode dat dit woord werd geoefend de woordenlijst werd afgedekt. Er werd ook een controlegroep opgenomen.

De 'woord vormen'- en 'probleem noemen'-taken werden op dezelfde manier uitge-

voerd als in Experiment 1, met dit verschil dat voor de kinderen in de 'probleem noemen'-conditie zowel de woorden als de target-letters of letterclusters nu op één bladzijde waren afgedrukt (in Experiment 1 stonden de woorden op een apart kaartje). De kinderen in de controleconditie verrichtten geen enkele trainingstaak. Bij hen werd alleen het dictee afgenomen.

Proefpersonen

Uit een nieuwe steekproef van 109 leerlingen uit Groep 3 van basisscholen die gebruik maken van de leesmethode "Veilig leren lezen" werd, op dezelfde manier als in Experiment 1, een groep van 72 kinderen opgenomen in de analyses. De groep proefpersonen uit Experiment 1 was dus ongelijk aan die van Experiment 2. Ook nu werd ervoor gezorgd dat in elke conditie het aantal proefpersonen gelijk was (n = 12). Net als in Experiment 1 werd de gemiddelde Z-score van de lees- en spellingscore vastgesteld en werd ervoor gezorgd dat de gemiddelde Z-scores tussen condities niet significant verschillend van elkaar waren ($F(5,66)$ = .011, p = 1.0). De gemiddelde scores waren 18.21 (SD 1.7) op de spellingtest en 44.64 (SD 14.4) op de leestest. De significante correlatie tussen de lees- en spellingtest bedroeg .41 ($p \leq$ .01). Het vooronderzoek vond ook nu plaats één week voor het experiment.

3.2 Resultaten

De bespreking van de resultaten valt opnieuw uiteen in de onderdelen trainingsfase en testfase. De fouten op het dictee werden op dezelfde wijze beoordeeld als in Experiment 1. Eerst zullen de resultaten van zowel de trainings- als testfase van Experiment 2 aan de orde komen. Daarna volgt een tweede analyse waarin de gegevens van Experiment 1 en de gegevens op het dictee van de makkelijke conditie van Experiment 2 zijn samengevoegd. Dit werd gedaan om het effect van tijdstip van testafname na te gaan.

Trainingsfase

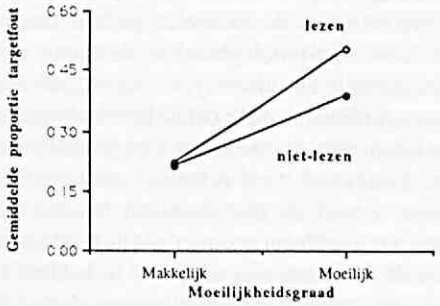
Uit een 2 (niet-lees-conditie: woord vormen vs. probleem noemen) bij 2 (moeilijkheidsgraad: makkelijk vs. moeilijk) bij 2 (taak: lezen vs. niet-lezen) variantie-analyse op de afhankelijke variabele tijdsduur met herhaalde metingen op de variabele taak bleek opnieuw een hoofdeffect van taak ($F(1,44)$ = 1776.48, p < .001). Het uitvoeren van de leestaak nam significant minder tijd in beslag (.54 min) dan het uitvoeren van de niet-lees-taak (5.53 min). Net als in Experiment 1 was ook hier het hoofdeffect van niet-lees-conditie significant ($F(1,44)$ = 391.06, p < .001). De 'probleem noemen'-conditie nam een significant kortere tijd in beslag (1.52 min) dan de 'woord vormen'-conditie (4.55 min). Een significant interactie-effect tussen niet-lees-conditie en moeilijkheidsgraad ($F(1,44)$ = 12.93, p < .001) gaf aan dat de gemiddelde tijd nodig om de 'woord vormen'-conditie in de moeilijke conditie uit te voeren significant hoger lag dan de tijd die hiervoor nodig was in de makkelijke conditie ($F(3,44)$ = 18.86, p < .001). In de 'probleem noemen'-conditie was geen verschil tussen de makkelijke en de moeilijke condities op de variabele trainingsduur. Tussen taak en moeilijkheidsgraad werd eveneens een significant interactie-effect gevonden ($F(1,44)$ = 49.24, p < .001). Over de leestaak werd in de moeilijke conditie gemiddeld .67 min gedaan en in de makkelijke conditie .42 min. Het uitvoeren van een niet-leestaak nam in de moeilijke conditie gemiddeld 6.49 minuten in beslag terwijl daar in de makkelijke conditie 4.57 minuten voor nodig waren.

Een variantie-analyse op het aantal fouten tijdens de training gaf aan dat er geen statistisch significant verschil bestond tussen de 'woord vormen'- en de 'probleem noemen' conditie. Wel werden tijdens de niet-leestaak (1.69) meer fouten gemaakt dan tijdens de leestaak (.56), $F(1,128)$ = 29.8, p < .001.

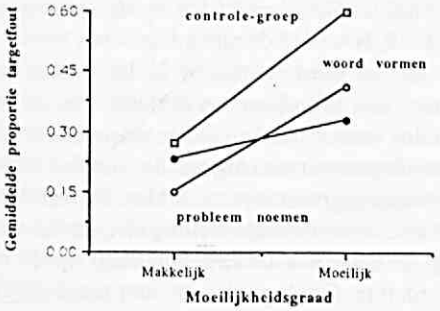
Testfase

Uit een 3 (niet-lees-conditie: woord vormen vs. probleem noemen vs. controle-groep) bij 2 (moeilijkheidsgraad: makkelijk vs. moeilijk) bij 2 (taak: lezen vs. niet-lezen) variantie-analyse op het aantal targetfouten met herhaalde metingen op de variabele taak bleken alle drie de hoofdeffecten significant. In de leestaak (.36) werden significant meer targetfouten gemaakt dan in de niet-lees-taak (.30), $F(1,66)$ = 7.48, p < .01. Het hoofdeffect van moeilijkheidsgraad was $F(1,66)$ = 33.14, p < .001. Het gemiddeld aantal fouten in de moeilijke conditie lag significant hoger (.44) dan het aantal in de makkelijke conditie (.21). Het

hoofdeffect van niet-lees-taak was ook significant, $F(2,66) = 6.57, p < .01$. Een post hoc analyse (Newman-Keuls, $p < .01$) maakte duidelijk dat er door de kinderen uit de 'woord vormen'- en 'probleem noemen'-condities (beide .28) significant minder targetfouten werden gemaakt dan door de kinderen in de controle-groep (.43). Er was een marginaal significante interactie tussen niet-lees-conditie en moeilijkheidsgraad, $F(2,66) = 2.84$, $.05 < p < .10$. In de makkelijke conditie werden door de kinderen in de 'woord vormen'- (.23), 'probleem noemen'- (.15) en controle-groep (.27) evenveel targetfouten gemaakt ($p > .20$), terwijl in de moeilijke conditie het aantal targetfouten in de 'woord vormen'- (.33) en 'probleem noemen'-groep (.41) significant ($F(2,66) = 7.93, p < .01$) lager was dan dat van de controle-groep (.60). De interactie tussen taak en moeilijkheidsgraad bleek significant te zijn ($F(1,66) = 3.96, p = .05$): in de moeilijke conditie was er wel een effect van taak ($F(1,66) = 11.15, p < .001$) maar in de makkelijke conditie niet. Beide interacties zijn weergegeven in de Figuren 1 en 2.

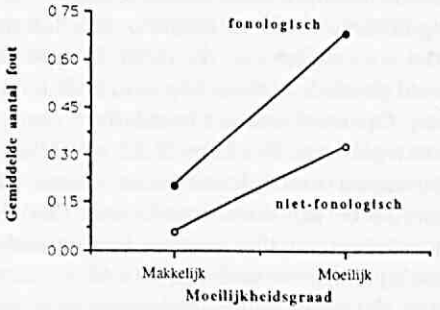


Figuur 1. Targetfouten in de lees- en niet-lees-condities als functie van de moeilijksheidsgraad van woorden (Exp. 2)



Figuur 2. Targetfouten in de niet-lees-conditie als functie van de moeilijkheidsgraad van woorden (Exp. 2)

Evenals in Experiment 1 werd er gekeken naar typen fouten die op de spellingtest werden gemaakt. Om na te gaan wat het effect van type fout was, werd eenzelfde analyse als hierboven uitgevoerd, maar met toevoeging van de variabele type fout. Er zal niet ingegaan worden op de hierboven reeds besproken hoofd- en interactie-effecten. Het effect van type fout bleek significant te zijn, $F(1,66) = 61.19, p < .001$. Het gemiddeld aantal fonologische fouten (.44) lag significant hoger dan het aantal niet-fonologische fouten (.19). Het interactie-effect tussen moeilijkheidsgraad en type fout was eveneens significant ($F(1,66) = 11.15, p < .01$). In de makkelijke conditie werden gemiddeld .20 fonologische en .06 niet-fonologische fouten gemaakt, in de moeilijke conditie bedroeg het gemiddeld aantal fonologische en niet-fonologische fouten .68 respectievelijk .33. Deze interactie is grafisch weergegeven in Figuur 3. De variabele type fout interacteerde met geen enkele andere variabele. Gemiddelde target-, fonologische en niet-fonologische fouten zijn weergegeven in de Tabellen 1 en 2 in Bijlage D.



Figuur 3. Fonologische en niet-fonologische fouten als functie van de moeilijkheidsgraad van woorden (Exp. 2)

3.3 Resultaten van de makkelijke conditie van Experimenten 1 en 2 gecombineerd

Trainingsfase

Een 2 (niet-lees-conditie: woord vormen vs. probleem noemen) bij 2 (tijdstip: februari vs. mei) bij 2 (taak: lezen vs. niet-lezen) variantie-analyse op de tijdsduur van de training met herhaalde metingen op de variabele taak liet een significant hoofdeffect van tijdstip zien ($F(1,48) = 9.09, p < .01$). In de maand mei nam

het uitvoeren van de training significant minder tijd in beslag (gemiddeld 2.58 min) dan in de maand februari (3.21 min). Het hoofdeffect van niet-lees-conditie was net als in de vorige analyses significant, $F(1,48) = 99.00, p < .001$. De gemiddelde trainingduur in de 'probleem noemen'-conditie was significant korter (1.85 min) dan in de 'woord vormen'-conditie (3.93 min). Het verschil tussen de tijdsduur van de leestaak (.89 min) en die van de niet-lees-taak (4.89 min) was ook significant, $F(1,48) = 1044.66, p < .001$. De significante interactie tussen tijdstip en niet-lees-conditie ($F(1,48) = 4.35, p < .05$) liet zien dat de tijdsduur voor alle niveaus verschilt tussen de tijdstippen van testafname (Newman-Keuls, $p < .01$), behalve bij het uitvoeren van de 'woord vormen'-conditie. Deze neemt in februari gemiddeld evenveel tijd in beslag als in mei.

Testfase

Een 2 (niet-lees-conditie: woord vormen vs. probleem noemen) bij 2 (tijdstip: februari vs. mei) bij 2 (taak: lezen vs. niet-lezen) variantie-analyse op het aantal targetfouten met herhaalde metingen op de variabele taak liet geen significante hoofd- en interactie-effecten zien. Met toevoeging van de variabele type fout werd dezelfde analyse uitgevoerd als hierboven. Opnieuw was het hoofdeffect van type fout significant ($F(1,48) = 22.22, p < .001$). Het gemiddeld aantal fonologische fouten (.20) overtrof het aantal niet-fonologische (.08). Het hoofdeffect van taak bereikte in deze analyse een bijna significant niveau ($F(1,48) = 3.72, p < .10$). Het gemiddelde aantal fouten op de leestaak (.16) was hoger dan op de niet-lees-taak (.12). De interactie tussen niet-leesconditie en type fout was niet significant. Analyses op gemiddelde target-, fonologische en niet-fonologische fouten zijn weergegeven in de Tabellen 1 en 2 in Bijlage D.

3.4 Conclusie

De resultaten van Experiment 2 komen overeen met die van Van Doorn-van Eijsden (1984) en Bosman en De Groot (1992). Ook hier is gebleken dat het leren van de spelling van een woord met behulp van een leestaak minder effectief is dan met een niet-leestaak (probleem noemen en woord vormen). Tussen de taken 'probleem noemen' en 'woord vormen' bleken geen verschillen te zijn op het dictee. De prestaties van de controle-groep waren beduidend slechter dan die van de experimentele groepen, en ook bleken makkelijke woorden beter gespeld te worden dan moeilijke woorden.

Het uitblijven van een differentieel effect van spellinginstructie-methoden in Experiment 1 is zeer waarschijnlijk toe te schrijven aan het feit dat de daar gebruikte stimuli te makkelijk waren. Uit de resultaten van Experiment 2 bleek inderdaad een interactie tussen de moeilijkheidsgraad van woorden en de variabele niet-lees-conditie. Kinderen in de makkelijke conditie (overeenkomstig met de situatie van Experiment 1) die met de 'probleem noemen'- of de 'woord vormen'-methode woorden leerden spellen vertoonden geen betere prestaties dan de kinderen in de controle-groep, terwijl in de moeilijke conditie er wel degelijk verschillen waren tussen de groepen.

Net als in Experiment 1 werden ook hier meer fonologische dan niet-fonologische fouten gemaakt. Dit resultaat steunt opnieuw de veronderstelling dat beginnende spellers na een spellinginstructie-training overwegend verklankend spellen. Uit het ontbreken van een interactie tussen de variabele type fout (fonologisch vs. niet-fonologisch) en de factor niet-lees-conditie en tussen type fout en taak mag geconcludeerd worden dat de hier onderzochte spellinginstructie-methoden niet geleid hebben tot kwalitatief verschillende spellingstrategieën. Zowel de niet-leestaken 'woord vormen' en 'probleem noemen' als de leestaak lijken de kinderen niet aangezet te hebben tot lexicaal spellen. De alternatieve verklaring dat de kinderen nog niet toe waren aan een strategie van lexicaal spellen, omdat hun spellingniveau daarvoor niet voldoende is, lijkt onwaarschijnlijk. Het gemiddelde op de spellingtest was 18.21 terwijl de maximum score 20 is.

Eerder werd verondersteld dat 'woord vormen' een grondiger verwerking van de stimulus vereist dan 'probleem noemen', en dat een diepere verwerking van de stimulus tot betere spellingprestaties zou leiden. De resultaten steunen deze veronderstelling niet, omdat noch op de targetfout-beoordeling noch op de beoordeling fonologische en niet-fonologische fouten een verschil was tussen de spelling-instructie-methoden 'woord vormen' en 'probleem noemen'. Toch willen wij hieruit niet

onmiddellijk de conclusie trekken dat diepte van verwerking geen invloed heeft op de volledigheid van de representatie. Een andere mogelijkheid is dat de assumptie dat 'woord vormen' een grondiger verwerking van de stimulus vereist dan 'probleem noemen' foutief is. Hoewel bij 'woord vormen' in tegenstelling tot 'probleem noemen', het hele woord gereconstrueerd moest worden, vergde dit geen bijzonder actieve houding, omdat de letters van het woord immers zichtbaar aanwezig waren. De proefpersoon moest met behulp van de blokjes de juiste volgorde bepalen van de letters waaruit het woord bestond. Alleen wat de target betreft moest een keuze gemaakt worden, omdat daarvoor een alternatief aanwezig was. De operationalisatie van de variabele diepte van verwerking is daarom mogelijk ongelukkig gekozen.

Uit het gegeven dat de 'probleem noemen'-training minder tijd kost dan de 'woord vormen'-training, maar dat beide even effectieve spellinginstructie-methoden zijn (zowel op de target- als de type fout beoordeling was er geen verschil tussen condities) kan geconcludeerd worden dat de 'probleem noemen'-instructie efficiënter is. De conclusie dat niet de duur, maar de kwaliteit van de bestede tijd doorslaggevend is voor de spellingresulaten lijkt gerechtvaardigd. Deze conclusie wordt ondersteund door de gecombineerde gegevens van Experiment 1 en 2. In februari hadden de kinderen in de 'probleem noemen'-conditie meer tijd nodig om de lees- en niet-leestaak uit te voeren dan in mei. De kinderen in de 'woord vormen'-conditie hadden in februari ook meer tijd nodig om de leestaak te voltooien dan die in mei, maar om de niet-leestaak uit te voeren hadden de kinderen in beide maanden evenveel tijd nodig. Dat de kinderen die in mei deelnamen aan de 'woord vormen' taak net zoveel tijd nodig hadden als de kinderen die in februari participeerden heeft zeer waarschijnlijk te maken met het feit dat het grootste deel van de 'woord vormen' taak in beslag genomen werd door het opzoeken, uitleggen en verwijderen van de letterblokjes. De aandacht in de 'woord vormen'-conditie is dus voor een groot deel niet-spelling gericht.

4 Algemene discussie

De hierboven gerapporteerde experimenten tonen aan dat onervaren spellers vooral verklankend spellen. Dat dit niet beperkt blijft tot de groep moeilijke woorden wordt ondersteund door de resultaten van het eerste experiment, met makkelijke woorden. Een mogelijk punt van kritiek op de door ons uitgevoerde experimenten is dat een gering aantal confrontaties met een moeilijk woord nooit aanleiding zou kunnen geven tot een lexicale strategie. Het feit echter dat ook de makkelijke woorden nog verklankend gespeld werden ondervangt dit punt.

Assink (1981) concludeerde op grond van zijn onderzoek met grammaticaal bepaalde homofonen (gebeurt of gebeurd) dat onervaren spellers (leerlingen afkomstig uit het Lager Beroeps Onderwijs, 16-18 jarigen) voornamelijk een woordbeeld- of lexicale strategie hanteren bij het bepalen van de spelling van werkwoordsvormen. Deze conclusie wordt hier niet aangevochten. Een kanttekening is wel dat de keuze van deze spellers voor een andere dan de lexicale strategie alleen dan mogelijk zou zijn als zij de grammaticale regels voor het bepalen van de juiste schrijfwijze zouden beheersen. Omdat de gemiddelde score van de onderzochte groep praktisch op kansniveau lag (51% correct) is het zeer onwaarschijnlijk dat zij inderdaad over die kennis beschikten. Deze veronderstelling wordt ondersteund door een onderzoek van Bosman en De Groot (in voorbereiding). Bij universitaire studenten (ervaren spellers) die de door Assink gebruikte taak uitvoerden waren er geen aanwijzingen dat de goede spellers vooral een lexicale strategie volgden. Daarentegen bleken de zwakke spellers wel overwegend lexicaal te spellen. Een aannemelijke verklaring hiervoor is dat ook deze groep zwakke spellers de regels niet goed beheerste en dus wel gedwongen werd om de lexicale strategie toe te passen.

Leren spellen blijkt niet simpelweg een kwestie van het herhaaldelijk lezen van het woord. Zoals in het onderzoek van Bosman en De Groot (1992) en Van Doorn-van Eijsden (1984) bleek ook nu dat lezen eigenlijk de minst geschikte manier is om te leren spellen. De opmerking in de handleiding van de methode "Veilig leren lezen", waarin gesteld wordt dat lezen alleen niet *altijd* voldoende is

om de spelling van een woord te leren, lijkt met weglating van het woord 'altijd' een betere weergave van het spelleerproces te zijn. De spellinginstructie-methoden: 'overschrijven', 'probleem noemen' en 'mondeling spellen' (in Bosman & De Groot, 1992) en 'woord vormen' en 'probleem noemen' uit Experiment 2 blijken betere spellingresultaten op te leveren dan lezen. De opmerking van Reitsma (1985, p. 82) dat je spellen vooral leert door te spellen lijkt gezien de resultaten van ons onderzoek volledig gerechtvaardigd.

Lezen en 'overschrijven' zijn de meest gebruikte didactische methoden om te leren spellen. Hoewel 'overschrijven' tot betere resultaten leidt dan lezen is deze methode niet bijzonder efficiënt. Zij vergt namelijk veel meer tijd dan bijvoorbeeld 'probleem noemen', maar heeft hetzelfde leereffect. De conclusie dat niet de duur, maar de kwaliteit van de geïnvesteerde tijd kritisch is lijkt gerechtvaardigd.

Uit het voorafgaande kan afgeleid worden dat de relatie tussen lezen en spellen minder hecht is dan aanvankelijk werd aangenomen (zie ook de inleiding). Deze conclusie wordt ondersteund door twee andere empirische gegevens. Op de eerste plaats blijkt de correlatie tussen lees- en spellingscores niet bijzonder hoog te zijn (in het eerste experiment .34 en in het tweede .41). Volgens Malmquist (1958) varieert deze tussen .50 en .80 (Frith, 1980). Ten tweede blijkt het veel gemakkelijker goede lezers te vinden die tegelijkertijd zwakke spellers zijn dan zwakke lezers die goed spellen. Beide bevindingen zijn moeilijk verenigbaar met de opvatting dat lezen en spellen omkeerbare processen zijn.

Samenvattend kunnen we zeggen dat beginnende lezers en spellers in alle onderzochte situaties een duidelijke voorkeur (blijven) vertonen voor een verklankende spelwijze en dat zij geen goede spellers worden door vaak te lezen.

Noten

1. De weergave van dit klassieke visuele woordherkenningsmodel is zeer summier, en bovendien is het model recentelijk niet van kritiek verschoond gebleven (ondere andere: Van Orden, Johnston & Hale, 1988).
2. Aan het eind van Groep 3 beheersen de kinderen de volgende categorieën: KM-, MK-, MKM-, MKMM-, MMKM-, MMKMM-, MMMKM-, en MKMMM- woorden. Daarnaast kunnen ze een beperkt aantal grondwoorden schrijven met een bepaalde spellingmoeilijkheid, waaronder de toonloze e of schwa (in bijvoorbeeld 'de'), sch-woorden, uitgangen -oor, eer, en uur, finale d (in be*d*), finale b (in he*b*), ei/ij, ch/g en au/ou onderscheid, -ng (in di*ng*) en -nk (in de*nk*) woorden, woorden met de uitgangen -eeuw (in sn*eeuw*), -oei (in ro*ei*), aai (m*aai*), -ieuw (in n*ieuw*), en open lettergrepen zoals -a, -o en -u (in ja, zo en nu).
3. De hier gebruikte woorden zijn allen unaniemen, dat wil zeggen dat 90% van een groot aantal Nederlandse en Vlaamse leerkrachten van mening is dat de betrokken woorden door 6-jarigen passief gekend moeten worden.
4. Experiment 1 werd uitgevoerd in een fase van het lees- en spelonderwijs waarin de kinderen alleen nog maar geconfronteerd zijn met het ambigue /ij/-foneem, de /au/- en de finale d als /t/-klank is dan nog niet onderwezen. Daarom is naast het /ij/-foneem gekozen voor de minder ambigue /s/-/z/ en /v/-/f/ als probleem fonemen. Het is bekend dat beginnende spellers (met name in het Westen van het land) soms problemen hebben met de /s/-/z/ en /v/-/f/. Uit een telling van het aantal woorden (totaal ca. 5550) uit de eerste drie leesboekjes van Veilig leren lezen is gebleken dat aan het begin van een woord 268 keer een woord met een 'Z' voorkomt, 45 met een 'F' en 263 met een 'V'. Daarnaast zijn er 228 woorden met een 'IJ' en 56 met een 'EI'. De ervaring met de onderzochte targets verschilt dus per letter.
5. De hier gebruikte spellingtest heeft een redelijk goede betrouwbaarheid (α = .88; n=218) en correleert r = .85 (n=207) met de spellingtest van Mommers en Van Dongen (1986).
6. Het beoordelen van spellingfouten op het al dan niet fonologisch zijn blijkt niet volledig systematisch te kunnen. De categorisering van fouten is gebaseerd op een aantal arbitrair gekozen regels waarbij de nadruk voornamelijk lag op interne consistentie. Een voorbeeld maakt het probleem duidelijk.

   Het incorrect gespelde woord 'blouwt' (blauw) werd als volgt beoordeeld: 1 fonologische fout ('ou' in plaats van 'au') en 1 niet-fonologische fout (toevoegen van de 't'). Daarentegen werden in het fout gespelde woord 'ijskauwd' de 'ou'/'au' verwisseling en de toegevoegde 'w' als 2 fonolo-

gische fouten gecategoriseerd. Overeenkomstig de beoordeling van de spelling 'blouwt' zou men ook kunnen redeneren: er is een letter toegevoegd dus 1 niet-fonologische fout en 1 fonologische fout vanwege de 'ou'/'au' verwisseling. Maar hier werd de voorkeur gegeven aan 2 fonologische fouten omdat de 'w' in een /ou/-cluster hoorbaar is.

Literatuur

Assink, E. M. H. (1981). Schrijfstrategieën van intuïtieve spellers bij moeilijke werkwoordsvormen. *Tijdschrift voor Taalbeheersing, 3*, 55-66.

Assink, E. M. H., Doorn-van Eijsden, M. van, & Hees, H. van (1982). Schrijfstrategieën van onervaren spellers bij lexicaal bepaalde homofonen. *Tijdschrift voor Taalbeheersing, 4*, 314-327.

Baron, J., Treiman, R., Wilf, J. F., & Kellman, P. (1980). Spelling and reading by rules. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 159-194). London: Academic Press.

Bosman, A. M. T., & Groot, A. M. B. de (1991). De ontwikkeling van woordbeelden bij beginnende lezers en spellers. *Pedagogische Studiën 68*, 199-215.

Bosman, A. M. T., & Groot, A. M. B. de (1992). Differential effectiveness of reading and non-reading tasks in learning to spell. In F. Satow & B. Gatherer (Eds.), *Literacy without frontiers: Proceedings of the 7th European and 28th United Kingdom Reading Association Annual Conference at Heriot-Watt University, Edinburgh, July 1991* (pp. 279-289). United Kingdom Reading Association, Widnes, Cheshire.

Caesar, F. B. (1970). *Eén-Minuut-Test voor de technische leesvaardigheid.* Tilburg: Zwijsen.

Caesar, F. B. (1979). *Veilig leren lezen: Structuurmethode voor het aanvankelijk leesonderwijs.* Tilburg: Zwijsen.

Caesar, F. B. (1980). Handleiding bij *Veilig leren lezen: Structuurmethode voor het aanvankelijk leesonderwijs.* Tilburg: Zwijsen.

Doorn-van Eijsden, M. van (1984). Leer je spellen door veel te lezen? *Tijdschrift voor Taalbeheersing, 6*, 252-263.

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

Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 495-515). London: Academic Press.

Jorm, A. J. (1981). Children with reading and spelling retardation: functioning of whole-word and correspondence-rule mechanisms. *Journal of Child Psychology and Psychiatry, 22*, 171-178.

Kohnstamm, G. A., Schaerlaekens, A. M., Vries, A. K. de, Akkerhuis, G. W., & Froonincksx, M. (1981). *Nieuwe streeflijst woordenschat voor 6-jarigen.* Lisse: Swets & Zeitlinger.

Kreiner, D. S. (1992). Reaction time measures of spelling: testing a two-strategy model of skilled spelling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 765-776.

Mommers, M., & Dongen, D. van (1986). *Zorgverbreding bij het leren lezen.* Tilburg: Zwijsen.

Nelson, H. E., & Warrington, E. K. (1974). Developmental spelling retardation and its relation to other cognitive abilities. *British Journal of Psychology, 65*, 265-274.

Reitsma, P. (1985). De verwerving van orthografische kennis. In E. Assink & G. Verhoeven (Eds.), *Visies op spelling* (pp. 73-82). Groningen: Wolters-Noordhoff.

Sloboda, J. A. (1980). Visual imagery and individual differences in spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 231-248). London: Academic Press.

Tenney, Y. J. (1980). Visual factors in spelling. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 215-229). London: Academic Press.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 371-386.

Waters, G. S., Bruck, M., & Seidenberg, M. (1985). Do children use similar processes to read and spell words? *Journal of Experimental Child Psychology, 39*, 511-530.

Manuscript aanvaard 11-1-1993

Auteurs

A. M. T. Bosman is als a.i.o. werkzaam bij de vakgroep Psychonomie van de Universiteit van Amsterdam.

M. van Leerdam is student bij de vakgroep Psychonomie van de Universiteit van Amsterdam.

Adres: Universiteit van Amsterdam, Faculteit der Psychologie, Vakgroep Psychonomie, Roetersstraat 15, 1018 WB Amsterdam

Abstract

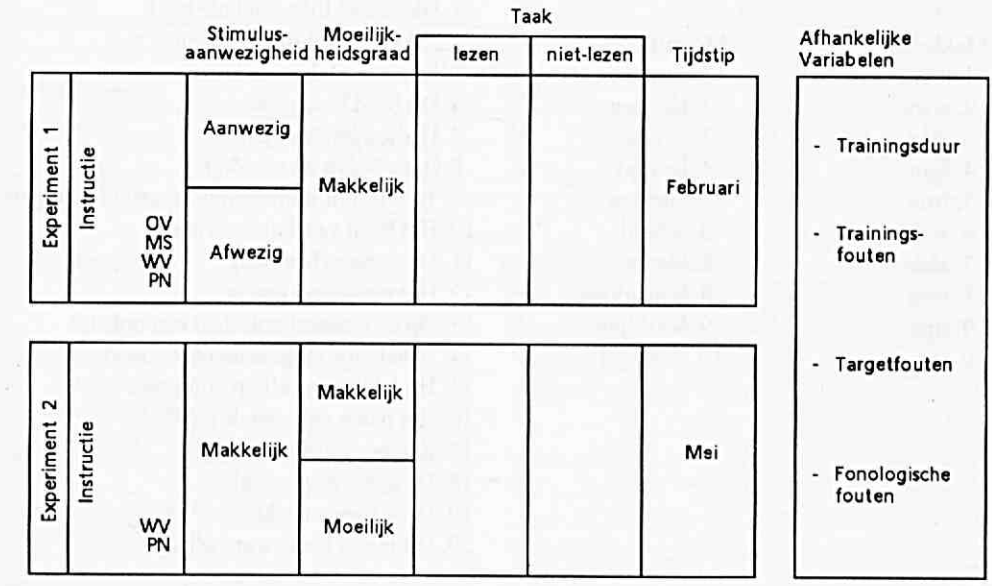
Beginning spelling: prevalence of the phonological strategy in spelling and the scant effectivity of reading as a spelling instruction method

A. M. T. Bosman & M. van Leerdam. Pedagogische Studiën, 1993, 70, 28-45.

This article focused on spelling instruction methods and spelling processes in beginning spellers (students from Grade 1). After a short training, the students predominantly used a phonological way of spelling and did not convert to a lexical strategy. This conclusion was based on the results of two experiments, in which the subjects made more phonological than non-phonological errors. Of the spelling instruction methods studied here reading appeared to be the least effective one. The two non-reading instruction methods 'problem naming' and 'word composition' had the same effects on spelling performance, but were both superior to reading. 'Problem naming' appeared to be more efficient than 'word composition', because it took longer for the subjects in the 'word composition' condition to acquire the same spelling level as those in the 'problem naming' condition.

Bijlage A

Experimentele procedure van Experiment 1 en 2



OV = overschrijven
MS = mondeling spellen
WV = woord vormen
PN = probleem noemen

Bijlage B

Stimulusmateriaal van de Experimenten 1 en 2

In Experiment 1 bestond het stimulusmateriaal uit de makkelijke woorden en in Experiment 2 werden zowel de makkelijke als de moeilijke woorden gebruikt. Van elk woord is de target onderstreept.

Makkelijk	Moeilijk
1. reis	1. strand
2. zoen	2. kauwen
3. vlag	3. ijskoud
4. lijm	4. bedtijd
5. fruit	5. zuchten
6. zeil	6. schuld
7. ziek	7. blauw
8. vies	8. bouwdoos
9. rijst	9. hoofdpijn
10. fee	10. stoplicht

Bijlage C

Spellingtest

De onderstreepte woorden vormden de testwoorden.

1. Er zit een gat in mijn broek.
2. Brrr..... wat een kou.
3. Die stapel lijkt een hele boel.
4. s'Avonds schijnt de maan.
5. Er zit een lek in de kraan.
6. Dit bord is schoon.
7. Die jongen heet Guus.
8. In de boom zit een duif.
9. Ik buk om de pen van de grond te rapen.
10. Het been van Jan deed zeer.
11. Deze zin is heel lang.
12. Het meisje eet een ei.
13. Op de vensterbank staat een pot.
14. Ik heb vijf vingers aan elke hand.
15. Hij heeft een bril op zijn neus.
16. Het pakje valt van de stoel.
17. Een poes lust graag melk.
18. De agent zegt: stop!
19. Deze banaan is krom.
20. Dit is een heel zware asbak.

Bijlage D

Tabel 1

*Gemiddeld aantal targetfouten in Exp. 1 (1), Exp. 2 (2), en Exp. 1 en 2-gecombineerd (1,2)**

		Stimulusaanwezigheid							
		Aanwezig				Afwezig			
Spellinginstructie		Lezen		Niet-lezen		Lezen		Niet-lezen	
(1)	OV	.32	(.18)	.20	(.21)	.25	(.20)	.26	(.24)
	MS	.20	(.17)	.34	(.25)	.33	(.16)	.30	(.15)
	WV	.20	(.20)	.18	(.19)	.31	(.16)	.27	(.12)
	PN	.31	(.24)	.26	(.20)	.31	(.21)	.27	(.22)
		Moeilijkheidsgraad							
		Makkelijk				Moeilijk			
Spellinginstructie		Lezen		Niet-lezen		Lezen		Niet-lezen	
(2)	WV	.23	(.22)	.22	(.18)	.42	(.25)	.23	(.19)
	PN	.17	(.17)	.13	(.13)	.47	(.13)	.35	(.19)
	CTR	.27	(.18)	.27	(.18)	.60	(.23)	.58	(.23)
		Tijdstip							
		Februari				Mei			
Spellinginstructie		Lezen		Niet-lezen		Lezen		Niet-lezen	
(1,2)	WV	.22	(.15)	.12	(.10)	.22	(.22)	.20	(.18)
	PN	.29	(.22)	.25	(.23)	.20	(.21)	.15	(.15)

Standaardafwijkingen staan tussen haakjes

OV = overschrijven
MS = mondeling spellen
PN = probleem noemen
WV = woord vormen

* minimum foutscore = 0
maximum foutscore = 1

Tabel 2

*Gemiddeld aantal type fouten in Exp. 1 (1), Exp. 2 (2), en Exp. 1 en 2-gecombineerd (1,2)**

		Type fout			
Spellinginstructie		Fonologisch		Niet-Fonologisch	
(1)	OV	.24	(.15)	.09	(.17)
	MS	.27	(.16)	.09	(.12)
	WV	.28	(.18)	.09	(.16)
	PN	.22	(.14)	.09	(.14)
(2)	WV	.39	(.29)	.14	(.16)
	PN	.38	(.33)	.21	(.22)
	CTR	.55	(.39)	.23	(.38)
(1,2)	WV	.19	(.15)	.08	(.13)
	PN	.22	(.17)	.08	(.11)

Standaardafwijkingen staan tussen haakjes

OV = overschrijven
MS = mondeling spellen
PN = probleem noemen
WV = woord vormen
CTR = controle-groep

* minimum foutscore = 0
maximum foutscore is onbepaald

Morele oordeelsvorming bij 'antisociale' jongeren en scholieren

J. R. P. B. de Mey

Samenvatting

Dit artikel doet verslag van een empirisch onderzoek naar de constructen moreel oordeelsstadium en het moreel inhoudelijke oordeel bij drie groepen jongeren: jeugdige delinquenten (*n*=60, gemiddelde leeftijd 16.2), leerlingen afkomstig van scholen voor Zeer Moeilijk Opvoedbare Kinderen (ZMOK'ers, *n*=150, gemiddelde leeftijd van 15.5) en leerlingen afkomstig uit het Middelbaar Economisch en Administratief Onderwijs (MEAO'ers, *n*=46, gemiddelde leeftijd 16.6). Zij scoorden op vraagstukken uit Kohlbergs Heinzdilemma en een zogenaamd Zorgonderdeel. Onderzocht is of de constructen discrimineren tussen de verschillende groepen en met elkaar samenhangen. De uitkomsten wijzen uit dat de jeugdige delinquenten en ZMOK'ers dezelfde morele interpretatiekaders hanteren en in dit opzicht verschillen van MEAO'ers, die op een hoger niveau redeneren. De moreel inhoudelijke oordelen geven alleen in de groep delinquenten en ZMOK'ers een relatie te zien met het algemene morele oordeelsniveau.

1 Inleiding

Dit artikel rapporteert over een empirisch onderzoek naar de morele oordeelsvorming bij een groep jeugdige delinquenten, een groep Zeer Moeilijk Opvoedbare Kinderen en een groep scholieren uit het Middelbaar Economisch en Administratief Onderwijs. Twee constructen van morele oordeelsvorming, die door Kohlberg en zijn collega's in de loop der tijd zijn uitgewerkt en relevant worden geacht voor de bestudering van moreel gedrag, krijgen hier nader aandacht, te weten het morele oordeelsstadium en het moreel inhoudelijke oordeel.

In eerdere studies beweert Kohlberg (1976, 1978) dat niet het inhoudelijke oordeel met betrekking tot goed of slecht in verband staat met moreel gedrag, maar het oordeelsstadium. Een oordeelsstadium is in Kohlbergs theorie een set criteria op grond waarvan het individu in morele dilemma's een inhoudelijk oordeel over morele waarden weet te maken, bijvoorbeeld een oordeel over leven of wet. Hij (Kohlberg, 1981) onderscheidt zes morele oordeelsstadia, die over drie niveaus verdeeld zijn. Deze laatste duidt hij aan als het preconventionele niveau (straf- en instrumentele gerichtheid), het conventionele niveau (oriëntatie op de groep en loyaliteit aan de wet) en het postconventionele niveau (oriëntatie op universele en ethische principes). De morele niveaus en stadia zijn hiërarchisch geordend, hetgeen betekent dat de hogere niveaus (en stadia) van een hoger moreel gehalte zijn dan de lagere. Ondersteund door empirisch materiaal, verzameld bij drie groepen jeugdige gevangenen (± 16 jaar), beweert hij, dat jeugdige delinquenten overwegend op een (lager) preconventioneel niveau oordelen. Individuen uit deze groep interpreteren goed en kwaad in termen van voordelige of nadelige gevolgen (straf, beloning, gunsten). Dit in tegenstelling tot een meerderheid 'normale' leeftijdgenoten, die op een (hoger) conventioneel niveau redeneren. Individuen uit deze groep interpreteren goed en kwaad in termen van groepsnormen. De resultaten van andere onderzoekers die deze hypothese onderzochten, zijn echter niet even eenduidig (Binder, 1988). Jammer is voorts, dat Kohlberg niet onderzocht heeft of zijn hypothese ook voor groepen geldt, waarvan het antisociale gedrag niet officieel ontdekt en bestraft is. Ook zou het bijvoorbeeld kunnen dat detentie effect heeft op de morele visie van gevangenen. Goffman (1968) beschrijft in dit verband gedetailleerd hoe 'inmates' onder institutionele dwang hun eigen morele codes ontwikkelen.

Na de delinquentenstudies (Kohlberg 1958, 1978) komen Kohlberg en zijn collega's (Jen-

nings, Kilkenny, Kohlberg, 1983; Kohlberg & Candee, 1984) terug op de uitspraak, dat het inhoudelijke oordeel er niet toe doet. Ook al is er een relatie tussen oordeelsstadia en moreel gedrag vastgesteld, onduidelijk blijft hoe en waarom moreel redeneren samenhangt met moreel gedrag. Men meent nu, daarbij Blasi (1983) volgend, dat voor de bestudering van moreel gedrag het inhoudelijke oordeel dat er direct aan voorafgaat, onontbeerlijk is. Ervan uitgaande dat hogere oordeelsstadia prescriptiever en universeler zijn dan lagere, suggereren Kohlberg en zijn collega's dat in de hogere oordeelsstadia meer individuen zich conform hun moreel inhoudelijke oordelen gedragen dan individuen in lagere stadia. Opvallend zou hierbij zijn dat in de hogere stadia meer individuen consensus hebben over inhoudelijke oordelen in een moreel dilemma dan individuen in lagere oordeelsstadia. Kohlberg en zijn collega's hebben zelf in dit verband geen empirisch onderzoek gedaan. Ze onderbouwen hun veronderstellingen met gegevens uit een studie van McNamee (1978) en een niet gepubliceerde studie van Helkema (1979). De gegevens zijn echter beperkt van aard. Op de eerste plaats beperken de onderzoeken zich slechts tot één morele situatie, zodat het de vraag is in hoeverre de resultaten generaliseerbaar zijn naar andere morele situaties. Onderzoek heeft uitgewezen dat individuen afhankelijk van de morele situatie vanuit een hoger of lager stadium kunnen redeneren (De Mey, 1991). Op de tweede plaats kunnen de resultaten vertekend zijn door het feit dat de gegevens bij studentenpopulaties verzameld zijn. Het gedrag van deze groepen, dat in relatie met de voornoemde morele oordeelsaspecten onderzocht is, is niet crimineel of antisociaal te noemen. Hierdoor is het de vraag of de gesuggereerde relaties tussen oordeelsstadia en inhoudelijke oordelen zich voordoen bij groepen, die antisociale gedragskenmerken hebben.

1.1 Probleemstelling en hypothesen

Naar aanleiding van de voorgaande beschouwingen luidt de probleemstelling in dit onderzoek als volgt: in hoeverre verschillen 'antisociale' groepen jongeren van elkaar en van 'niet-antisociale' leeftijdgenoten in hun vermogen tot moreel oordelen? Als zich in dit opzicht verschillen aftekenen, in hoeverre is er dan overeenstemming in inhoudelijke oordelen waar te nemen? Het probleem is onderzocht bij drie groepen jongeren (± 16 jaar): jeugdige delinquenten in strafinrichtingen, Zeer Moeilijk Opvoedbare Kinderen (voortaan ZMOK'ers) en leerlingen uit het Middelbaar Economisch en Administratief Onderwijs (voortaan MEAO'ers). Hoewel jeugdige delinquenten en ZMOK'ers van elkaar te onderscheiden zijn, wijzen verschillende onderzoekers (Kazdin, 1987; Witte, 1986) op veel overeenkomsten tussen beide groepen, anderen scheren beide groepen zonder meer over één kam (Gibbs 1987). De overeenkomsten zijn zodanig dat op beide groepen de term antisociaal van toepassing is (Kazdin[1], 1987). De term antisociaal gedrag is hier gebruikt als een brede aanduiding van handelingen die op onrechtmatige wijze aan anderen fysieke, geestelijke of materiële schade toebrengen.

Het verschijnsel 'delinquentie' betreft gedrag waarbij strafbare feiten gepleegd zijn en waarvoor een veroordeling door de rechter heeft plaatsgevonden. In de context van dit artikel is dat gevangenisstraf. ZMOK'ers zijn jongeren die door belangrijke anderen (ouders, onderwijzers) aangemerkt zijn als antisociaal, maar niet juridisch vervolgd zijn. Het zijn scholieren uit het Voortgezet Speciaal Onderwijs. Dit soort onderwijs volgen kinderen en jongeren die agressief zijn, zich niet aan de gangbare regels willen en kunnen houden en zeer beperkte sociale vaardigheden hebben. Het is gedrag dat de basale behoeften van anderen frustreert alsook de rechten van anderen met voeten treedt. Deze groep is hier interessant, omdat onderzoek naar hun morele ontwikkeling antwoord geeft op de vraag in hoeverre zij in dit opzicht van jeugdige delinquenten verschillen.

Gelet op de in de literatuur genoemde overeenkomsten, is de verwachting dat jeugdige delinquenten en ZMOK'ers niet verschillen in hun oordeelsstadia *(hypothese 1a)* en volgens Kohlberg op preconventioneel niveau oordelen *(hypothese 1b)*. De groep MEAO'ers, die als 'normaal' beschouwd is, dient ter vergelijking met de andere twee groepen. De verwachting is dat MEAO'ers hoger scoren dan de twee andere groepen en wel op conventioneel niveau *(hypothese 1c)*. Naar aanleiding van de veronderstelling, dat er een (monotone) relatie is tus-

sen oordeelsstadia en inhoudelijke oordelen, zijn de volgende verwachtingen onderzocht: a) jeugdige delinquenten laten in hun inhoudelijke oordelen evenveel overeenkomst zien als ZMOK 'ers (*hypothese 2a*), b) MEAO'ers vertonen meer overeenkomst in hun inhoudelijke oordelen dan de twee andere groepen (*hypothese 2b*) en c) in de betreffende groepen geven individuen in hogere oordeelsstadia meer overeenkomst in hun inhoudelijke oordelen te zien dan individuen in lagere oordeelsstadia (*hypothese 2c*).

2 Methode

2.1 Onderzochte groepen

De jeugdige delinquenten (n=60) met een gemiddelde leeftijd van 16.2 (minimum 14 jaar, maximum 18 jaar) zaten ten tijde van het onderzoek in twee Amsterdamse strafinrichtingen. De leerlingen van scholen voor Zeer Moeilijk Opvoedbare Kinderen waren afkomstig uit Amsterdam, Haarlem, Utrecht, Rotterdam, Middelburg (n=150), met een gemiddelde leeftijd van 15.5 (minimum 14 jaar, maximum 18 jaar). De leerlingen afkomstig uit het Middelbaar Economisch en Administratief Onderwijs, afkomstig uit Hoorn (n=46), hebben een gemiddelde leeftijd van 16.6 (minimum 16 jaar, maximum 18 jaar).

Hoewel het leeftijdsverschil tussen de groepen niet groot is, is gezocht naar eventuele leeftijdseffecten. De ZMOK'ers zijn daarom verdeeld in twee groepen, een groep met gemiddelde leeftijd 16.5 (n=33) en een groep met gemiddelde leeftijd 14.1 (n=56). Vergelijking van deze twee groepen op de algemene morele niveauindex (zie 2.4) leverde geen significant verschil (t=1.56, df=87, p=.123) op.

Gebruikmakend van de beroepenclassificatie van het CBS (1984) is de sociale achtergrond van de respondenten onderzocht. De jeugdige delinquenten , de ZMOK'ers en de MEAO'ers komen in het algemeen uit de lager sociaal-economische milieus[2].

2.2 Instrumentarium

Gebruikt is het bekende Heinzdilemma uit de Sociomoral Reflection Objective Measure (SROM) (vgl. Gibbs, Arnold, Morgan, Schwartz, Gavaghan & Tappan, 1984) en een destijds door ons ontworpen instrument, het Zorgonderdeel (De Mey, Van der Draai & Spiecker, 1989).

In het *Heinzdilemma* verkeert Heinz in de situatie, dat zijn vrouw komt te sterven als zij niet geholpen wordt met een net ontdekt medicijn. Heinz heeft niet genoeg geld voor het medicijn en de apotheker weigert het medicijn goedkoop aan Heinz te verstrekken. Wat moet Heinz doen? De wet overtreden door het medicijn te stelen, waarmee z'n vrouw gered wordt, of het naderende sterven van z'n vrouw aan te zien en dus de wet te respecteren? In dit dilemma zijn metingen gedaan op inhoudelijke oordelen en oordeelstadia.

De inhoudelijke oordeelsvariabelen. De respondent moet hierop aangeven welk gedrag hij in een bepaalde situatie moreel juist vindt. Bijvoorbeeld, moet Heinz stelen of moet hij niet stelen op de vraag: 'Wat als de stervende persoon niet Heinz' vrouw is, maar een vriend. Wat zou Heinz moeten doen?'

De stadiumvariabelen. De respondent moet hierop aangeven welke voorgegeven zes antwoordmogelijkheden in de buurt komen van zijn antwoord op een bepaald vraagstuk over het Heinzdilemma, waarna hij aan moet geven welk antwoord het meest in de buurt komt. De voorgegeven antwoordmogelijkheden corresponderen met Kohlbergs oordeelsstadia, een zesde antwoordmogelijkheid is een zogenaamd pseudo-, een 'onzin'-alternatief dat een controlefunctie heeft voor het wel of niet serieus beantwoorden van de vragen door de respondent (zie ook procedure). Een voorbeeld moge het een en ander verduidelijken. De respondent krijgt een variant op het Heinzdilemma voorgelegd, waarin de stervende persoon niet Heinz' vrouw maar zijn vriend is. De respondent kan de volgende redenen aankruisen: a) omdat je vriend misschien dingen voor jou heeft gedaan zou je hem een dienst moeten bewijzen, zeker als je wilt dat je vriend je in de toekomst ook helpt, b) omdat een vriendschap gebaseerd moet zijn op wederzijds respect en samenwerking, c) omdat het je vriend is die een beroemd persoon kan zijn, d) omdat je je betrokken zou voelen bij je vriend en zou verwachten dat je vriend jou ook zou helpen, e) omdat jij en je vriend misschien volledig aan elkaar toegewijd zouden zijn geraakt, f) (*pseudo*) omdat een eerste vereiste van gene-

genheid een relatie is.

In het *Zorgonderdeel* staan variabelen centraal, die betrekking hebben op situaties, waarin de respondent de relatie met anderen geproblematiseerd ziet. De situaties geven een discrepantie aan tussen het belang van de respondent en de relatie met anderen, en doen een beroep op de zorg en verantwoordelijkheid van de respondent. In dit onderdeel zijn om redenen van praktische aard alleen de oordeelstadia van de respondenten gemeten, die naar analogie van de voorgegeven antwoordmogelijkheden uit het Heinzdilemma geformuleerd zijn. Zo wordt bijvoorbeeld beschreven dat de respondent met zijn/haar vrienden/vriendinnen in de disco is, waar het pas laat heel dol wordt. Hoewel de ouders er nooit wat van zullen zeggen, weet de respondent dat zij ongerust zijn bij laat thuiskomen. De respondent kan de volgende redenen aanvoeren om op tijd thuis te komen: a) omdat je straf krijgt als je niet doet wat je ouders zeggen, b) omdat, als jij geen rekening houdt met je ouders, je kunt verwachten dat zij voortaan geen rekening meer houden met jouw wensen, c) omdat het niet fair (sportief) is om te laat thuis te komen als je weet dat je ouders ongerust zijn, d) omdat je je in redelijkheid naar de wensen van je ouders hoort te schikken, e) omdat het in een samenleving noodzakelijk is dat de jongere generatie de oudere generatie respecteert, f) (*pseudo*) omdat ongerustheid van ouders beloftes noodzakelijk maakt.

2.3 Procedure

De gegevens zijn verzameld met behulp van klassikaal afgenomen vragenlijsten. Van te voren is met een aantal groepsleiders en docenten de vragenlijst doorgenomen en op enkele plaatsen bijgesteld. Aangezien verwacht kon worden dat een aantal respondenten uit de groepen jeugdige delinquenten en ZMOK'ers niet over al te grote leesvaardigheden beschikte, is groepsleiders en docenten verzocht de vraagstukken voor te lezen en zo nodig toelichting te geven, uiteraard zonder in te gaan op de mogelijke antwoordalternatieven. Voor iedere groep is onderzocht in hoeverre de verzamelde protocollen wel of niet voor verdere analyse in aanmerking kwamen. Een protocol met zeven of meer pseudo's deed niet mee, omdat dan aangenomen werd dat de rest van de ingevulde vragen onbetrouwbaar was (Gibbs et al., 1984). Op grond van deze analyse is de onderzoeksgroep delinquenten gereduceerd tot n=45, ZMOK'ers tot n=89, MEAO'ers tot n=40.

2.4 Scoring en analyse

De scores op de inhoudelijke oordelen behoefden geen bewerkingen te ondergaan. De scores op de stadiumitems zijn volgens een procedure van Gibbs et al. (1984) bewerkt tot itemindexen en een algemene morele niveauindex[3]. Op basis van deze indexen is voor individuen respectievelijk het morele oordeelsniveau per onderwerp en het morele oordeelsniveau in het algemeen vast te stellen. Hierbij betekent een hogere score een meer gevorderde morele ontwikkeling. De scores zijn geanalyseerd met variantieanalyses (MANOVA, SPSS/PC+) en meervoudige vergelijkingstoetsen (Duncans test, SPSS/PC+), op de gelijkheid van gemiddelden, tweezijdig op 5%- niveau. Naar verschillen in inhoudelijke overeenstemming tussen en in de groepen is gezocht met χ^2-toetsen.

2.5 Metrische eigenschappen

Het omschreven instrumentarium vertoont voor eerder onderzochte groepen gunstige metrische eigenschappen[4]. De Mey et al. (1989) rapporteerden destijds voor MEAO'ers en ZMOK'ers Cronbachs alpha's van .70 en .71. De betrouwbaarheid voor de hier onderzochte groepen, gemeten over 32 items met Cronbachs alpha, resulteerde in het volgende. Voor de groep jeugdige delinquenten en ZMOK'ers is een alpha van .75 genoteerd. Voor de ZMOK'ers, een beduidend grotere groep dan de groep die in 1989 is onderzocht, is een alpha van .76 gevonden en voor de delinquenten is dat een alpha van .71. Voor de MEAO'ers, een geheel andere groep dan die in 1989 is bestudeerd, is een alpha van .76 genoteerd.

3 Resultaten

3.1 Vergelijking van jeugdige delinquenten, ZMOK'ers en MEAO'ers op oordeelsstadia

Voor de toetsing van de hypothesen 1a, 1b en 1c, die betrekking hebben op stadiumverschillen tussen de groepen, zijn multivariate variantieanalyses (MANOVA, SPSS/PC+) en post-hoc-toetsen uitgevoerd. Eerst is een variantieanalyse uitgevoerd op de algemene morele ni-

Tabel 1

Gemiddelden ($\bar{x}$), standaardafwijkingen (sd) voor delinquenten, ZMOK'ers en MEAO'ers, univariate F-waarden (2, 171) voor groepseffecten

	Delinquenten		ZMOK'ers		MEAO'ers		
item indexen	$\bar{x}$	*sd*	$\bar{x}$	*sd*	$\bar{x}$	*sd*	*F*-waarde
Heinz1	259	(166)	253	(154)	336	(146)	4.17*
Heinz2	247	(128)	241	(151)	322	(99)	5.28**
Heinz3	151	(137)	155	(141)	205	(152)	1.99
Heinz4	243	(153)	244	(160)	310	(126)	2.99
Heinz5	158	(138)	177	(139)	270	(141)	8.16***
Heinz6	156	(154)	180	(154)	287	(134)	9.44***
Heinz7	197	(136)	197	(135)	313	(83)	13.34***
Heinz8	181	(138)	216	(148)	204	(144)	.87
Heinz9	219	(173)	194	(161)	218	(137)	.52
Heinz10	214	(134)	165	(155)	291	(159)	9.64***
Zorg11	140	(182)	157	(173)	274	(179)	7.51***
Zorg12	255	(146)	181	(144)	325	(122)	15.57***
Zorg13	208	(182)	157	(173)	214	(173)	2.04
Zorg14	217	(167)	236	(172)	319	(132)	4.87**
Zorg15	258	(189)	216	(207)	353	(188)	6.59***
Zorg16	222	(136)	207	(154)	256	(128)	1.63
Hotellings T^2= .64							3.11***
SROMTOT	208	(59)	199	(65)	281	(53)	26.65***

* : $p < .05$
** : $p < .01$
***: $p < .005$

Tabel 2

Resultaten van Duncans multipele range toetsen[5] op gemiddelde scores (tussen haakjes) van delinquenten (A), ZMOK'ers (B) en MEAO'ers (C) en F-waarden (2, 171)

Itemindexen	Duncans vergelijking	*F*-waarde
Heinz1	A (259) = B (253) < C (336)	4.17
Heinz2	A (247) = B (241) < C (321)	5.28
Heinz5	A (157) = B (177) < C (270)	8.16
Heinz6	A (156) = B (180) < C (287)	9.44
Heinz7	A (196) = B (197) < C (313)	13.34
Heinz10	A (213) = B (165) < C (290)	9.64
Zorg11	A (140) = B (157) < C (273)	7.51
Zorg12	A (255) > B (181) < C (326)	15.57
Zorg14	A (217) = B (236) < C (319)	4.87
Zorg15	A (258) = B (216) < C (353)	6.59
SROMTOT	A (208) = B (199) < C (281)	26.65

veauindex, gecontroleerd voor leeftijd. Dit resulteerde in een significant groepseffect {F (2, 170)= 27.54, $p< .001$}. Dit was aanleiding om verder te zoeken naar groepseffecten op specifieke itemindexen. In Tabel 1 is te zien, dat het multivariate testresultaat voor de afhankelijke itemindexen een significant groepseffect {Hotellings T^2= .64, F (32, 310)= 3.11, $p< .001$} aangeeft.

De univariate testresultaten in Tabel 1 wijzen groepseffecten aan voor het merendeel van de itemindexen. Om te bepalen tussen welke groepen verschillen optreden zijn post-hoc-toetsen uitgevoerd op de itemdexen, waar significante groepseffecten voor geregistreerd zijn. De resultaten staan in Tabel 2.

In Tabel 2 is te zien dat met uitzondering van Zorg12 (voorbeelden van de items staan in 2.2) tussen jeugdige delinquenten en ZMOK'ers geen verschillen optreden. Op deze itemindex scoren jeugdige delinquenten hoger dan ZMOK'ers. Deze uitkomst is niet zo verwonderlijk. Het item vraagt redenen waarom het belangrijk is om niet mee te doen aan de klein-

ste diefstallen. Het is niet ondenkbaar dat jeugdige delinquenten anders dan ZMOK'ers over het moreel onjuiste van dit soort vergrijpen zijn gaan denken door de argumenten die de rechter heeft aangevoerd bij het uitspreken van de veroordeling. Op grond van deze uitkomsten kan worden volgehouden dat jeugdige delinquenten dezelfde morele interpretatiekaders hanteren als ZMOK'ers (hypothese 1a). Gelet op de gescoorde oordeelsstadia, de indeling van Gibbs et al. (1984) in aanmerking nemend, is in Tabel 1 ook te zien dat beide groepen in het algemeen in Kohlbergs tweede oordeelsstadium redeneren, dus niet boven het preconventionele niveau uitkomen, waarmee Kolbergs bevindingen (hypothese 1b) ook voor deze groepen gelden.

In de Tabellen 1 en met name 2 is te zien dat de vastgestelde groepseffecten vooral voor rekening van de MEAO'ers komen. Zowel op de algemene morele niveauindex als op de gemarkeerde itemindexen is te zien, dat MEAO'ers significant hoger scoren. Voorts is te zien, de indeling van Gibbs et al. (1984) in aanmerking nemend, dat de morele niveauindex op conventioneel niveau ligt. Conform dit resultaat kan hypothese 1c niet verworpen worden.

3.2 Vergelijking van jeugdige delinquenten, ZMOK'ers en MEAO'ers en 'lage' en 'hoge' scoorders in de groepen op overeenstemming over inhoudelijke oordelen

Voor de toetsing van de hypothesen 2a en 2b, dat jeugdige delinquenten over hun inhoudelijke oordelen evenveel overeenstemming hebben als ZMOK'ers, maar MEAO'ers in dit opzicht van beide groepen verschillen, zijn χ^2 -toetsen uitgevoerd. De resultaten staan in Tabel 3. Kortheidshalve is volstaan met de scoringen op één antwoordalternatief. (Het andere alternatief is steeds het tegenstelde, bijvoorbeeld 'hij zou niet moeten stelen' en telt op tot honderd procent).

Tabel 3 laat zien, dat slechts op één oordeel (leven2) een groepseffect optreedt. Nadere bestudering van dit groepseffect leert dat jeugdige delinquenten en ZMOK'ers evenveel overeenstemming hebben op dit oordeel $\{\chi^2(1)= .54, p= .46\}$, maar dat MEAO'ers van delinquenten $\{\chi^2(1)= 4.61, p= .03\}$ en ZMOK'ers $\{\chi^2(1)= 8.24, p= .004\}$ verschillen. Dit betekent dat MEAO'ers eensgezinder zijn dan de twee andere groepen als het gaat om het oordeel, dat, om het leven van een vriend te redden, Heinz moet stelen. Concluderend is vast te stellen dat

Tabel 3

Percentages en resultaten van χ^2-toetsen op verschil in overeenstemming over de inhoudelijke oordelen tussen delinquenten, ZMOK'ers en MEAO'ers

Oordelen	Delinquent	ZMOK	MEAO	χ^2
leven 1				
Heinz moet stelen	78.4	69.7	87.5	4.05
leven 2				
Heinz moet stelen	60.6	52.6	87.0	8.23*
leven 3				
Heinz moet stelen	79.4	61.4	54.2	4.72
straf 1				
Heinz moet vrij	90.2	76.9	78.1	3.17
straf 2				
Heinz moet vrij	93.9	84.7	82.6	2.05
straf 3				
Heinz moet gevangen	81.1	82.6	91.2	1.65

* $p < .05$

Tabel 4

χ^2-toetsen op verschillen tussen 'lage'en 'hoge' scoorders voor overeenstemming over inhoudelijke oordelen. Delinquenten en ZMOK'ers vormen één groep

Moreel oordeel	'Antisociaal' laag	hoog	χ^2	MEAO'ers laag	hoog	χ^2
leven-items						
H. moet stelen	68.3	77.4	1.15	89.5	84.6	.17
H. moet stelen	60	51.1	.72	84.6	90.0	.15
H. moet stelen	66.7	68	.02	56.3	50.0	.08
straf-items						
H. moet vrij	72.5	90.9	6.06*	82.4	73.3	.38
H. moet vrij	80	95.7	5.41*	81.8	83.3	.01
H. gevangen	71.9	93.9	8.63**	94.7	86.7	.68

* : $p < .05$
**: $p < .01$

er tussen jeugdige delinquenten en ZMOK'ers geen verschil in overeenstemming op inhoudelijke oordelen is, hetgeen voor handhaving van hypothese 2a pleit. Maar hypothese 2b, MEAO'ers vertonen meer overeenkomst in hun inhoudelijke oordelen dan de jeugdige delinquenten en ZMOK'ers, moet worden verworpen. Het contrast tussen de groepen komt, zoals vermeld op slechts één oordeel naar voren.

Toetsing van hypothese 2c, dat in de betreffende groepen individuen in hogere oordeelsstadia meer overeenkomst in hun inhoudelijke oordelen te zien geven dan individuen in lagere oordeelsstadia, is uitgevoerd met behulp van χ^2-toetsen. De groepen jeugdige delinquenten en ZMOK'ers zijn nu voor de vergelijking met MEAO'ers als één groep behandeld. Tussen de eerste twee groepen is immers geen verschil in oordeelsstadium en ook geen verschil in overeenstemming over inhoudelijke oordelen waargenomen. De respondenten in de groepen zijn verdeeld over twee categorieën, 'laag' en 'hoog', waarbij de mediaan van de morele niveauindex als grenswaarde genomen is (antisociaal: 200, MEAO'ers: 291). De resultaten staan in Tabel 4.

In Tabel 4 is te zien dat significante verschillen tussen 'hoge' en 'lage' scoorders alleen optreden bij jeugdige delinquenten en ZMOK'ers, dat deze verschillen voorkomen op de zogenaamde strafitems en dat er in de groep MEAO'ers geen verschillen optreden tussen 'hoge' en 'lage' scoorders. Op grond van deze resultaten kan niet zonder meer tot verwerping dan wel tot acceptatie van hypothese 2c worden overgegaan. Een nuancering van de hypothese is noodzakelijk. En wel als volgt. Naarmate een groep in het algemeen op een hoger moreel niveau redeneert wordt het verschil in inhoudelijke overeenstemming tussen 'lage' en 'hoge' scoorders kleiner.

4 Discussie en conclusies

Voor de hier onderzochte groepen moet worden vastgesteld, dat jeugdige delinquenten en ZMOK'ers dezelfde (preconventionele) morele interpretatiekaders hanteren, waarbij zij zich onderscheiden van de groep MEAO'ers, die op een significant hoger (conventioneel) moreel niveau redeneert. Concluderend kan worden vastgesteld dat dit resultaat bijdraagt aan de generalisatie van Kohlbergs bevindingen inzake zijn delinquentenstudies. Een andere verdienste is, dat deze studie één van de weinige onderzoeken is, die empirische gegevens aandraagt met betrekking tot de overeenkomsten in de morele ontwikkeling van jeugdige delinquenten en ZMOK'ers.

Voorzichtigheid is geboden met de acceptatie van Kohlbergs bewering dat een onder-

scheid in morele oordeelsniveaus tot uitdrukking komt in een verschil in overeenstemming over inhoudelijke oordelen. Kennelijk is het geconstateerde statistische verschil tussen enerzijds de jeugdige delinquenten en ZMOK'ers en anderzijds de MEAO'ers te gering voor een duidelijk contrast in inhoudelijke oordelen. Pas als in de groepen 'lage' van 'hoge' scoorders zijn onderscheiden, komt deze relatie tussen oordeelsstadium en inhoudelijke oordelen naar voren. Hierbij is opvallend, dat in de groep MEAO'ers 'hoge' en 'lage' scoorders inhoudelijk evenveel overeenstemming hebben, hetgeen vermoedelijk met hun algemeen hogere oordeelsniveau samenhangt. De zogenaamde leven- en strafitems zijn voor deze groep misschien te 'gemakkelijk'. Vooralsnog lijkt het raadzaam om de veronderstelling van Kohlberg en Candee (1984), dat er een relatie is tussen oordeelsstadium en inhoudelijk oordeel, niet te verwerpen en onderzoek hiernaar voort te zetten.

De vergelijking tussen enerzijds jeugdige delinquenten en ZMOK'ers en anderzijds MEAO'ers wekt misschien verwondering, omdat niet duidelijk is of en op welke manier het verschil in morele oordeelsvorming in relatie staat tot het morele gedrag van de groepen. Hier is geprobeerd aan de oplossing van dit probleem dat in de Kohlbergtraditie onderkend is een bijdrage te leveren door de relatie tussen morele oordeelsvorming en inhoudelijke oordelen te onderzoeken. De vergelijking die hiervoor plaatsvond tussen de groepen heeft vooral betrekking op hoog en laag scorende groepen.

Hoewel de vastgestelde overeenkomsten tussen jeugdige delinquenten en ZMOK'ers aansluiten bij de constateringen in de literatuur, is het onderscheid duidelijk. De ene groep is veroordeeld en zit in een strafinrichting en de andere groep gaat naar school. Ook al zijn beide groepen als 'antisociaal' aan te merken, de veronderstelling dat deze groepen verschillen in hun (a)morele gedrag is op basis van dit onderscheid aannemelijk. Geconstateerd moet dan worden dat de onderzochte morele oordeelsaspecten deze niet kunnen verklaren. Interessant is om na te gaan in hoeverre een ander moreel oordeelsaspect, door Blasi (1983) het verantwoordelijkheidsoordeel ('responsibility judgment') genoemd, een bijdrage kan leveren aan deze verklaring. Want volgens hem moet het moreel inhoudelijke oordeel voordat het tot een morele handeling leidt, door het 'filter' van het verantwoordelijkheidsoordeel.

Enige terughoudendheid in de uitspraken over jeugdige delinquenten en ZMOK'ers is misschien gepast, aangezien een aanzienlijk deel niet aan de vereiste scoringscriteria voldeed. Dat wil zeggen dat er in deze groepen veel respondenten waren, die meer dan de toegestane pseudo-antwoorden gaven (zie 2.3). Het is mogelijk dat hierdoor bij beide groepen een selectie heeft plaatsgevonden, waardoor de groepen niet representatief zijn voor de doorsnee jeugdige delinquent en ZMOK'er.

Tot besluit nog dit. In weinig studies naar de morele ontwikkeling van adolescenten wordt tevens naar een verband gezocht met moreel gedrag. Kohlberg en zijn collega's hebben hiervoor in theoretisch opzicht enkele bouwstenen aangedragen. Het lijkt er echter op, dat van de morele oordeelsconstructen, het oordeelsstadium en het inhoudelijk oordeel, vooral het eerste discriminerende en dus realiteitswaarde heeft. Maar dit construct alleen is niet voldoende om de brug naar moreel gedrag te slaan.

Noten

1 In de Engelstalige literatuur wordt de term 'conduct disorder' gebezigd, hetgeen wil zeggen "clinically severe antisocial behavior in which the everyday functioning is impaired as defined by parents, teachers, and others" (Kazdin 1987, p. 22). Uit de omschrijving is af te leiden dat het hiermee aangeduide gedragsverschijnsel hetzelfde is als het gedrag van ZMOK'ers.

2 Dat wil hier zeggen dat de vaders in beroepsgroepen horen, die aangeduid zijn als 'lagere employé's of 'arbeiders'.

3 De itemindexen zijn geconstrueerd met behulp van de handleiding van Gibbs et al. (1984). Hierbij zijn de volgende stappen gevolgd: 1) per itemonderdeel is een gemiddelde stadiumscore berekend over antwoorden die als 'in de buurt' zijn aangemerkt, 2) aan ieder itemonderdeel is een stadiumscore toegekend, gelet nu op het antwoord dat als het 'meest in de buurt' is aangekruist, 3) op 1 en 2 is een totaal gemiddelde berekend, waarbij de 'meest in de buurt'-scores twee maal zo zwaar wegen als de 'in de buurt'-scores.

De algemene morele niveauindex is een gewogen gemiddelde van de 'in de buurt'- en de 'meest in de buurt'-scores.

4 Vergelijking met de berekende alpha's van andere onderzoekers, zoals Gibbs et al.(α= .82, 1984) en Van IJzendoorn (α= .76, 1986), is niet goed mogelijk omdat hier maar ten dele, alleen het Heinzdilemma, gebruik is gemaakt van de door hen gebruikte SROM.

5 Behalve voor Zorg15 geven Duncans vergelijkingen dezelfde resultaten als Scheffé's testen. De Scheffé test voor Zorg15 geeft A = B, B < C, A = C.

Literatuur

Binder, A. (1988). Juvenile delinquency. *Annual Review of Psychology, 39*, 253-282.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin, 88*, 1-45.

Blasi, A.(1983). Moral cognition and moral action: A theoretical perspective. *Developmental Review, 3*, 178-210.

Centraal Bureau voor de Statistiek (1984). *Beroepenclassificatie*. Voorburg

Gibbs, J. C. (1987). Social processes in delinquency: The need to facilitate empathy as well as sociomoral reasoning. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Moral behavior through social interaction*. New York: Wiley.

Gibbs, J. C., Arnold, K. D., Morgan, R. L., Schwartz, E. S., Gavaghan, M. P., & Tappan, M. P. (1984). Construction and validation of a multiple-choice measure of moral reasoning: *Child Development, 55*, 527-536.

Goffman, E. (1968). *Asylums-essays on the social situation of mental patients and other inmates*. Great Britain: C. Nicholls & Company Ltd.

Helkema, K. (1979). *The development of the attribution of responsibility. A critical survey of empirical research and a theoretical outline*. Research Reports, Department of Social Psychology. Unpublished manuscript, University of Helsinki.

IJzendoorn, M. H. van (1986). Is Kohlbergs theorie van de morele ontwikkeling 'seksistisch'? *Kind en Adolecent, 7*, 34-39.

Jennings, W. S., Kilkenny, R., Kohlberg, L. (1983). Moral-development theory and practice for youthful and adult offenders. In W. S. Laufer & J. M. Day (Ed.), *Personality theory, moral development and criminal behavior*. Massachusets: Heath.

Kazdin, A. E. (1987). *Conduct disorders in childhood and adolescence*. London: Sage.

Kohlberg, L. (1958). *The development of modes of moral thinking and choice in the years 10 to 16*. Unpublished doctoral dissertation, University of Chicago.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive developmental approach. In T. Lickona (Ed.), *Moral development and behavior*. New York: Holt, Rinehart and Winston.

Kohlberg, L. (1978). The cognitive developmental approach to behavior disorders: A study of the development of moral reasoning in delinquents. In G. Serban (Ed.). *Cognitive defects in the development of mentall Ilness*. New York: Brunner/ Mazel.

Kohlberg, L. (1981). *The philosophy of moral development, moral stages and the idea of justice* (Vomule 1). San Francisco: Harper and Row Publishers.

Kohlberg, L., & Candee, D. (1984). The relationship of moral judgment to moral action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Morality, moral behavior and moral development*. New York: Wiley.

McNamee, S. (1978). Moral behavior, moral development and motivation. *Journal of Moral Education, 7*, 27-32.

Mey, J. R. P. B. de, Draai, J. van der, & Spiecker, B. (1989). Zorg en rechtvaardigheidsoordelen van scholieren, een empirisch onderzoek naar de controverse tussen Kohlberg en Gilligan. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs, 5*, 338-349.

Mey, J. R. P. B. de (1991). Diversiteit in de morele oordelen van scholieren: een onderzoek naar Kohlbergs structurele eenheidspostulaat. *Pedagogische studiën, 68*, 265-276.

Witte, H. F. J. de (1986). Psychopathoforme gedragsstoornissen en ontwikkelingspsychopathie. In J. A. R. Sanders-Woudstra & H. F. J. de Witte (Eds.), *Leerboek kinder-en jeugdpsychiatrie*. Assen: Van Gorcum.

Manuscript aanvaard 23-11-1992

Auteur

J. R. P. B. de Mey (1950) is als universitair docent verbonden aan de vakgroep Pedagogiek van de Vrije Universiteit.

Adres: Sectie Theoretische en Historische Pedagogiek, Transitorium I, Van der Boechorststraat 1, 1081 BT, Amsterdam.

Abstract

Moral judgments of 'antisocial' adolescents and students

J. R. P. B. de Mey. Pedagogische Studiën, 1993, 70, 46-55.

This article presents the results of an empirical study aiming to evaluate the constructs moral stage and moral choice. Three groups participated in this study: delinquent adolescents (n=60, mean age 16.2), students from several schools for Children with Severe Conduct Disorders (ZMOK students, n=150, mean age 15.5) and students from a College of Economic and Administrative Studies (MEAO students, n=46, mean age 16.6). They rated moral problems of one of Kohlberg's dilemma's, the Heinz dilemma, and a part with so-called Care-items. Analysis of variance (MANOVA, SPSS/PC+) and χ^2 tests are used to find out the relation between the constructs and their discriminating power for the groups just mentioned. The results show that delinquent adolescents and ZMOK students are reasoning in the same moral way in contrast to the MEAO students, who are reasoning at a higher level. Relations between the moral choices and the general moral reasoning stage only show up in the delinquent and ZMOK group.

Differentiële effecten van computergestuurde instructie op het lezen van dyslectische en zwak lezende kinderen: een voorstudie

H. Smeets en A. van der Leij*

Samenvatting

De invloed van computergestuurde instructie op de technische leesvaardigheid werd in deze voorstudie nagegaan bij 48 zwakke lezers met ongeveer een jaar achterstand en 77 dyslectische lezers met ruim twee jaar achterstand (resp. acht en bijna tien jaar oud). Twee vragen waren aan de orde: 1) Is het effect van intensieve computergestuurde instructie afhankelijk van de ernst van de leesproblemen (zwak versus dyslectisch)? 2) Is er een effect van intensieve computergestuurde instructie op de ontwikkeling (tot een geautomatiseerde vaardigheid) van lexicale en fonologische woordverwerking (als basisaspecten van technische leesvaardigheid)? De resultaten indiceren dat deze vorm van computergestuurde instructie bij beide groepen lezers effect sorteerde, maar meer bij de zwakke dan de dyslectische groep. Bij de zwakke lezers bracht de instructie een zeker inhaaleffect teweeg in het automatiseringsproces dat zich onder andere uitte in een betere verwerking van hoogfrequente en pseudo-woorden als operationalisaties van respectievelijk lexicale en fonologische verwerking. Bij de dyslectische lezers was de vooruitgang geringer. In de discussie worden de resultaten geïnterpreteerd onder andere in termen van een verschil tussen de twee groepen om tot automatisering van de twee verwerkingswijzen te komen.

Inleiding

In dit artikel wordt een vooronderzoek gerapporteerd waarin is nagegaan in hoeverre een computergestuurd trainingsprogramma effect heeft op de technische leesvaardigheid in het algemeen en de automatisering hiervan in het bijzonder bij kinderen met (ernstige) leesproblemen. Het gebruik van de computer als instrument om deze automatisering te bewerkstelligen, heeft voordelen die het traditioneel onderwijs niet kan bieden. De computer is een bij uitstek geschikt instrument voor tijdkritische stimuluspresentatie en registratie van de responstijd, twee elementen die een belangrijke rol spelen in het bewerkstelligen en het meten van automatisering (zie bijv. Van den Bosch, 1991). Bovendien is er de mogelijkheid tot verregaande individualisering van de instructie (Van Lieshout, 1990).

1 Theoretisch kader

In het leren lezen van woorden kunnen twee verwerkingswijzen worden onderscheiden die samen de pijlers vormen onder de automatisering van het leesproces. In het beginstadium leert de leerling woorden voornamelijk *fonologisch* te verwerken (fonologisch decoderen). Woorden worden via klank-tekenkoppeling en auditieve synthese in de klankvorm omgezet en verder verwerkt. Daarnaast ontwikkelt zich in de loop van het leerproces een tweede verwerkingswijze. Door oefening worden steeds meer woorden via *lexicale* verwerking (ook directe herkenning genoemd) gekoppeld aan de betekenis (Perfetti, 1985). Een lezer herkent woorden die hij al vele malen gezien heeft aan de orthografische informatie, zonder ze eerst via fonologisch decoderen om te hoeven zetten in de klankvorm. De globale vorm plus de analyse van enkele letters zijn voldoende om het woord te identificeren. Beide verwerkingswijzen worden in de loop van de eerste onderwijsjaren geautomatiseerd. Dit betekent dat de volleerde lezer in principe alle woorden snel en accuraat

* De auteurs danken E. J. Kappers en F. A. Goossens voor hun kritische commentaar op een eerdere versie van dit artikel.

kan verwerken, waarbij de snelheid een asymptotisch niveau bereikt (LaBerge & Samuels, 1974; Samuels, 1985). De fonologische verwerkingswijze gebruikt hij voor het lezen van onbekende, de lexicale voor bekende woorden. Door beheersing van de twee pijlers onder het automatiseringsproces van het technisch lezen, die zich uit in een toenemende accuratesse en snelheid, is de jonge lezer steeds beter in staat om zijn aandacht te richten op het begrijpen van de tekst (verg. Boland, 1991). Geautomatiseerde vaardigheden kosten immers nauwelijks aandacht (Schneider & Shiffrin, 1977).

Er zijn indicaties dat dyslectische kinderen, die gekenmerkt worden door een onverwachte discrepantie tussen hun gemiddelde intelligentieniveau en hun leesvaardigheid (Dumont, 1990), een ernstige achterstand hebben in de ontwikkeling van beide verwerkingswijzen (verg. Van der Leij, 1983; Spear & Sternberg, 1987). Daarbij is het, op grond van de wetenschappelijke evidentie, aannemelijk dat het probleem in de eerste verwerkingswijze, het fonologisch decoderen, als ernstiger en meer primair kan worden beschouwd (verg. bijv. Snowling, 1987; Rack, Snowling & Olson, 1992). Verschil van mening is er echter over de vraag in hoeverre dit probleem het gevolg is van een cognitieve ontwikkelingsachterstand (Treiman & Hirsh-Pasek, 1985) dan wel van een specifiek, dat wil zeggen structureelbelemmerend, tekort (Olson, Kliegl, Davidson & Foltz, 1985). Voor de instructiegevoeligheid van dyslectische kinderen maakt de aard van het onderliggende probleem verschil. Is er sprake van een ontwikkelingsachterstand dan loopt de leesontwikkeling weliswaar sterk achter, maar is hun leesvaardigheid inzake scoreprofielen op componentiële vaardigheden vergelijkbaar met die van jongere, normale lezers van hetzelfde leesniveau. Dientengevolge kan worden voorspeld dat deze kinderen uiteindelijk beide verwerkingswijzen zullen leren beheersen op automatisch niveau, zij het op een latere leeftijd en met veel extra oefening. Als daarentegen de oorzaak moet worden gezocht in een specifiek tekort, hetgeen tot uitdrukking komt in verschillende scoreprofielen, dan is het onwaarschijnlijk dat deze kinderen een of beide verwerkingswijzen naar een automatisch niveau kunnen tillen: zij zullen steeds verder achterraken.

Naar aanleiding van eerder uitgevoerd longitudinaal onderzoek (Van der Leij, in press; Yap, 1993) achten we het onwaarschijnlijk dat dyslectische kinderen woorden automatisch kunnen leren verwerken, zeker wanneer het gaat om woorden van een fonologisch en orthografisch complexer niveau dan de MKM-structuur (M=medeklinker, K=klinker), bijv. MMKM-woorden en twee-lettergrepige woorden. De hypothese van het specifieke tekort heeft derhalve onze voorkeur. Nog onduidelijk is echter of middels intensieve oefening de gestaag groeiende achterstand in de leesontwikkeling zou kunnen worden verminderd en eventueel weggewerkt. In het laatste geval zou een structureel tekort als verklaringsmodel minder waarschijnlijk worden.

Onderzoek toont aan dat intensieve oefening in de vorm van computergestuurde training gunstige effecten heeft op de leesvaardigheid van dyslectische kinderen (zie o.a. Frederiksen, Warren, & Roseberry, 1985a; 1985b; Van Daal, Van der Leij & Geervliet-van der Hart, 1989; Van den Bosch, 1991). Deze experimenten zijn echter in het algemeen te kleinschalig van opzet en te kort van duur om een uitspraak te kunnen doen over de vraag of het automatiseringsproces werkelijk op gang is gekomen.

Om meer duidelijkheid te verschaffen over de vraag of langdurige oefening bijdraagt tot de automatisering van de leesvaardigheid bij dyslectische kinderen, is een computergestuurd trainingsprogramma ontwikkeld op basis van de resultaten van een aantal deelexperimenten waarover elders is gerapporteerd (Van Daal, Van der Leij, Bakker en Reitsma, 1987; Van der Leij, 1990). In tegenstelling tot de oefeningen in deze deelexperimenten waarin slechts een beperkt aantal trainingscondities werd beproefd onder begeleiding van een proefleider, bestaat het programma dat in dit artikel wordt besproken uit een serie gevarieerde trainingstaken, die middels de computer volledig interactief, dus zonder proefleider, kunnen worden afgewerkt. Om na te gaan in hoeverre dit programma toepasbaar en effectief is, is een voorstudie opgezet. Daarin werd naast een groep dyslectische leerlingen, die ongeveer twee en een half jaar achterstand in leesontwikkeling had, gekozen voor een groep zwakke lezers met ruim één jaar achterstand. De vergelijking van

deze twee groepen aan de onderzijde van de normale leesontwikkelingscurve leek zowel uit theoretisch als praktisch oogpunt relevanter dan een vergelijking van meer extreme groepen als dyslectische en normale lezers. Theoretisch relevanter omdat uit eerder onderzoek (Yap & Van der Leij, in press) is gebleken dat zwakke lezers op een serie taken wel van dyslectici verschillen maar niet van normale lezers (zoals gematcht op leesniveau). Binnen de dichotomie van de achterstand versus het specifieke tekort lijken de zwakke lezers zich meer aan de zijde van de achterstand te bevinden, dat wil zeggen dat ze zich behalve door hun traagheid, verder niet door een abnormale leesontwikkeling profileren (verg. o.a. Stanovich, Nathan & Zolman, 1988). We veronderstelden dat, wanneer dyslectische lezers te lijden hebben onder een specifiek tekort en zwakke lezers niet, dit differentiële trainingseffecten tot gevolg zou hebben. Uit praktisch oogpunt was deze keuze relevanter omdat nagegaan kon worden in hoeverre het oefenprogramma toepasbaar en effectief is bij groepen lezers die verschillen in leesachterstand.

Op basis van de resultaten van de eerdere deelexperimenten verwachtten we dat computergestuurde training het lezen positief zou beïnvloeden. Onze voornaamste interesse was in hoeverre er sprake zou zijn van verschillende vormen van generalisatie van het geleerde. Langdurige training met een programma waarin verscheidene deelvaardigheden van het lezen geoefend worden (voor een beschrijving zie 3.3), leidde tot de verwachting dat die deelvaardigheden buiten het direct geleerde flexibel toegepast zouden worden. Ten eerste was de bijdrage aan de automatisering van beide verwerkingswijzen een centraal punt van aandacht. Wat betreft het fonologisch verwerken moest dit blijken uit het snel en accuraat verwerken van niet-geoefende pseudo-woorden (betekenisloze, maar volgens de regels van de Nederlandse taal correcte en uitspreekbare woorden) en wat betreft het lexicaal verwerken uit het snel en accuraat herkennen van niet-geoefende echte woorden. Deze pseudo- en echte woorden werden in de voor- en natoets computergestuurd en één voor één aangeboden, analoog aan de trainingssituatie. Ten tweede waren we geïnteresseerd in de vraag of eventuele trainingseffecten zouden generaliseren naar een betere woordverwerking in genormeerde leestaken waarin de woorden niet-computergestuurd en op een andere wijze werden aangeboden (rijtjes woorden; verhaaltjes). Ten derde kon een kwalitatief betere verwerking blijken uit generalisatie van oefening in het stillezen naar het hardop lezen van niet-geoefende woorden (het programma was interactief en deed dus geen beroep op het hardop lezen; de voor- en natoetsen daarentegen wel). Onderzoeksresultaten wijzen er op dat dyslectische leerlingen relatief slecht zijn in het hardop lezen, mogelijk als gevolg van problemen in het associëren van de orthografische structuur van een woord met de uitspraakvorm (Van der Leij, Smeets & Van Daal, 1990). Gerapporteerd is dat oefeningen waarbij het toetsenbord gebruikt wordt (spelling) even effectief zijn met betrekking tot hardop lezen als oefeningen in het hardop lezen, echter niet in de zin van generalisatie naar niet-geoefende woorden (Van Daal, Van der Leij & Geervliet-van der Hart, 1989). We verwachtten dat langduriger en gevarieerder oefening dit effect wel teweeg zou kunnen brengen.

2 Onderzoeksvragen

Deze overwegingen hebben geleid tot de volgende twee onderzoeksvragen:

1) Is het effect van intensieve computergestuurde instructie afhankelijk van de ernst van de leesproblemen? De verwachting hierbij is dat zwakke lezers meer zullen profiteren dan dyslectici. Indien het juist is dat dyslectici lijden onder een specifiek tekort en zwakke lezers niet, dan is het waarschijnlijk dat de laatste groep op de natoets hoger zal scoren dan de eerste.
2) Is er een effect van intensieve computergestuurde instructie op de ontwikkeling (tot een geautomatiseerde vaardigheid) van de lexicale en fonologische verwerking van woorden? Bij beide groepen werd exploratief nagegaan of er een effect is ten aanzien van betere, dat wil zeggen accuratere en snellere lexicale en fonologische woordverwerking (echte versus pseudo-woorden).

Voor de beantwoording van de eerste vraag is gebruik gemaakt van twee genormeerde toetsen. Daarnaast is een computergestuurde toets

afgenomen waarin woorden op diverse wijzen werden aangeboden en verwerkt (voor een beschrijving zie 3.2).

Ter beantwoording van de tweede vraag is m.b.t. de *accuratesse* gebruik gemaakt van een bij de voor- en natoets computergestuurd aangeboden hardop-leestaak, met als enige moeilijkheidsfactor de toenemende complexiteit van woordklassen (resp. MKM, MKMM, MMKM, twee-lettergrepige woorden en woorden met open lettergreep). In deze taak wordt het werkgeheugen nauwelijks belast, noch worden eisen gesteld omtrent een snelle verwerking van woorden (het woord blijft op het scherm staan). Een accuratere verwerking wordt geoperationaliseerd als het successievelijk toegenomen aantal beheerste woordklassen. De gebruikte klassen maken het mogelijk een eventueel effect te differentiëren naar fonologische en orthografische complexiteit. Bij één-lettergrepige woorden is sprake van toenemende fonologische complexiteit doordat er meer medeklinkers aan het begin of eind van een woord kunnen staan, terwijl de omzetting van orthografische naar klankvorm regelmatig blijft (resp. MKM versus MKMM en MMKM). Bij twee-lettergrepige woorden zijn er daarenboven verschillen in orthografische complexiteit doordat de juiste herkenning hier wel afhankelijk is van de orthografie (vergelijk bijvoorbeeld 'bommen' met 'bomen' waarin het aantal letters /m/ bepaalt of de klinker als een korte /o/ of als een lange /oo/ wordt uitgesproken). Door vergelijking van twee-lettergrepige woorden die bijvoorbeeld gesloten en/of open lettergrepen bevatten, kan de invloed van toenemende orthografische complexiteit worden nagegaan. Op grond van eerder uitgevoerd onderzoek (Van der Leij, in press) en de vigerende opvatting dat de kern van het probleem te kenschetsen is als een fonologisch tekort (o.a. Snowling, 1987) verwachtten we dat de dyslectische groep meer moeite zou hebben met toenemende fonologische complexiteit dan de zwak lezende groep. Er zou daarentegen geen differentieel effect zijn van toenemende orthografische complexiteit, los van de verwachting dat dyslectische kinderen met beide woordklassen meer moeite zouden hebben dan zwakke lezers.

De invloed van computergestuurde instructie op een mogelijk *snellere* lexicale en fonologische verwerking, is nagegaan door de resultaten van de twee groepen lezers op bovengenoemde hardop-leestaak, waarbij woorden een onbeperkte tijd worden aangeboden, in voor- en natoets te vergelijken met de resultaten op een hardop-leestaak waarbij woorden geflitst werden aangeboden: het effect van presentatietijdverkorting. De rationale daarachter is dat een leerling woorden die onbeperkt worden aangeboden, correct kan benoemen door ze om te zetten in de klankvorm. Echter wanneer de stimulus slechts korte tijd beschikbaar blijft, is het noodzakelijk om snel - en dus op zekere hoogte automatisch - te werken. Presentatietijdverkorting is in het onderzoek een veel gebruikt middel om de problemen van dyslectische kinderen met de snelheid van verwerking te duiden (Bouma & Legein, 1980; Van den Bosch, 1991).

3 Methode

3.1 Proefpersonen

Om bovenstaande onderzoeksvragen te kunnen beantwoorden hebben we twee groepen lezers geselecteerd: zwakke en dyslectische lezers. De groep zwakke lezers (ZWAK; n=48) was afkomstig uit de groepen drie, vier en vijf van twee basisscholen. Als criterium voor zwak lezen werd een achterstand van circa één jaar gehanteerd in het lezen van losse woorden, gemeten met behulp van de Eén-Minuut-Test (EMT, Brus & Voeten, 1973).

De selectie vond plaats op grond van de discrepantie tussen het aantal maanden genoten leesonderwijs en de bij de EMT gehanteerde gemiddelde normscores, die overeenkomen met het aantal maanden leesonderwijs waarop een bepaalde score gemiddeld gehaald wordt. Kinderen met een ruwe EMT-score <10 werden niet geselecteerd vanwege te verwachten problemen met de klank-tekenkoppeling. Leerlingen met een ruwe EMT-score >40 werden eveneens uitgesloten vanwege de veronderstelling dat ze reeds een bepaalde mate van automatisering bereikt hebben met betrekking tot de woordklassen die in de training en in voor- en natoets gebruikt werden.

De groep dyslectische lezers (DYS; n=77) was afkomstig van vier LOMscholen. Geselecteerd werden leerlingen met een achterstand

Tabel 1
Proefpersonen per groep

groep	n	j.	m.	leeftijd (mnd.) gem.	leeftijd (mnd.) sd	leesniveau[1] gem.	leesniveau[1] sd
DYS	77	63	14	117.8	13.1	11	4.0
ZWAK	48	29	19	100.9	10.5	13	4.2
totaal	125	92	33				

1. Leesniveau is gebaseerd op de score van de EMT en wordt uitgedrukt in didactische leeftijdsequivalent; 11 komt overeen met een gemiddeld leesniveau na één maand in groep vier, 13 correspondeert met een gemiddeld leesniveau na drie maanden in groep vier.

van tenminste twee jaar ten opzichte van het normale gemiddelde voor hun leeftijd. Ook hier gold dat hun EMT-scores dienden te liggen tussen 10 en 40. De EMT-scores van beide groepen verschilden niet significant van elkaar. Alle leerlingen lazen op een niveau dat gemiddeld gehaald wordt aan het begin van groep vier in het basisonderwijs. De groep dyslectici was echter bijna anderhalf jaar ouder dan de groep zwakke lezers.

De gegevens van de proefpersonen staan weergegeven in Tabel 1.

3.2 Tests

Onder gebruikmaking van een pretest-posttest design werden alle lezers vooraf en achteraf getoetst. De gebruikte toetsen waren de reeds vermelde EMT, de AVI-toetsen van Van den Berg en Te Lintelo (1977) en de Computergestuurde Toets voor Automatisering van het Lezen (CO-

Tabel 2
Vijf experimentele taken (COTAL)

A-V	*Auditief-Visueel matchen* Een auditieve stimulus (A: een hoogfrequent of pseudo-woord) wordt per koptelefoon gepresenteerd. Vervolgens verschijnt een visueel-orthografische stimulus op het scherm (V: hetzelfde of een ander woord). De V-stimulus blijft geprojecteerd totdat een antwoord is gegeven. De proefpersoon dient door middel van het klikken met de muis op een button aan te geven of beide stimuli hetzelfde zijn of verschillend. De moeilijkheidsgraad kan worden gevarieerd door het tijdsinterval tussen A- en V-stimulus in te stellen op 1, 3 of 5 seconden.
L	*Benoemen van woorden* Een visueel-orthografische stimulus (een hoogfrequent of pseudo-woord) wordt op het scherm gepresenteerd en de proefpersoon dient het woord hardop te lezen. Er zijn in deze taak geen moeilijkheidsniveaus: het woord blijft op het scherm totdat een antwoord is gegeven.
FL	*Benoemen van flitswoorden* Een visueel-orthografische stimulus (een hoogfrequent of pseudo-woord) wordt geflitst op het scherm aangeboden en onmiddellijk gemaskeerd m.b.v. een abstracte figuur. De proefpersoon wordt gevraagd het woord hardop te lezen. De moeilijkheidsgraad kan worden gevarieerd door de aanbiedingstijd van de woorden in te stellen op 200 of 100 msec.
V-T	Visueel Typen Een visueel-orthografische stimulus (een hoogfrequent of pseudo-woord) wordt geflitst op het scherm aangeboden en onmiddellijk gemaskeerd. De proefpersoon dient het woord nu na te typen. De moeilijkheidsgraad kan worden gevarieerd door de aanbiedingstijd van de woorden in te stellen op 200 of 100 msec.
A-T	Auditief Typen Een auditieve stimulus (A: een hoogfrequent of pseudo-woord) wordt per koptelefoon gepresenteerd. De proefpersoon dient het woord nu na te typen. Het moeilijkheidsniveau kan worden gevarieerd door de tijd tussen het aanbieden van de auditieve stimulus en de mogelijkheid tot antwoorden in te stellen op 1 of 3 seconden

TAL). COTAL is een toets die de beheersing van een aantal (deel)vaardigheden op het terrein van het technisch lezen tracht te meten. In vijf verschillende taken worden vier dimensies gevarieerd: a) aanbiedings- (auditieve of visueel-orthografische presentatie van woorden) en responsmodaliteit (matchen, typen of benoemen); b) moeilijkheidsgraad (geflitste aanbieding, interstimulustijd in matchingstaken); c) woordklasse (één- en twee-lettergrepige woorden); d) woordfrequentie (pseudo-woorden en hoogfrequente woorden). Voor een vollediger beschrijving zij verwezen naar Van der Leij, Smeets en Van Daal (1990). In het kader van dit onderzoek volstaan we met een korte beschrijving van de vijf taken (Zie Tabel 2). De hoogfrequente woorden waren afkomstig uit vier klassen: MKM-woorden, een combinatie van MKMM- en MMKM-woorden, twee-lettergrepige woorden en woorden met open lettergrepen. De pseudo-woorden bestonden uit drie éénlettergrepige klassen (MKM-, MKMM- en MMKM-woorden).

In iedere taak, woordklasse, moeilijkheidsgraad en versie (hoogfrequent of pseudo) werden per leerling 10 items (woorden of woordparen) at random geselecteerd uit een bestand. De kans dat woorden meer dan één keer werden aangeboden was relatief klein, zodat iedere afname binnen een woordklasse feitelijk een parallelversie betrof.

Testprocedure. De taken werden afgenomen onder leiding van een proefleider. Binnen een taak vond de afname plaats per woordklasse van eenvoudig naar moeilijk. Er werd een beheersingscriterium gehanteerd van 90% per taak. Beheersing van een taak resulteerde in een score van 1; werd het criterium niet gehaald dan was de score 0.

3.3 Het trainingsprogramma

Het Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen (COPAL), is een programma waarmee kinderen met leesproblemen zelfstandig en zonder tussenkomst van een leerkracht of proefleider lees- en spellingoefeningen op woordniveau kunnen doen (Van der Leij & Reitsma, 1990; Van der Leij, 1990). COPAL bestaat uit een aantal oefeningen die per leerling nader gespecificeerd en als pakket samengesteld kunnen worden (zie voor een overzicht Tabel 3). De leerlingen kunnen

Tabel 3
Onderdelen van COPAL

Oefening	Beschrijving
Structuurrij	Er verschijnen drie woorden op het scherm in de vorm van een structuurrij. Van een vierde woord bepaalt het kind of het wel of niet in het rijtje thuishoort.
Typen	Er verschijnt gedurende onbepaalde tijd een woord boven aan het scherm. Het kind typt dit woord over in een vak onder aan het scherm.
Visueel Dictee	Een woord verschijnt gedurende vijf seconden op het scherm. Vervolgens verdwijnt het woord en het kind typt dit dan na.
Flitsdictee	Een woord wordt geflitst aangeboden (1, 1/2, 1/3. 1/4, 1/5 sec.). Het kind typt het woord na.
Wat hoor je?	Er verschijnen vier woorden op het scherm. Een wordt er auditief aangeboden. Het kind kiest met de muis uit het rijtje het woord dat het heeft gehoord.
Wat zie je?	Een woord verschijnt op een scherm. Het kind kiest uit vier via de koptelefoon auditief aangeboden woorden welk het visueel gepresenteerde is.
Kies Teken bij Klank	Er verschijnt een MKM-woord op het scherm waarvan elke letter in een apart vak staat. Het kind krijgt een klank uit het woord te horen en bepaalt dan welke letter erbij hoort.
Vul Teken In	Er verschijnt een MKM-woord op het scherm waarvan een teken ontbreekt. Het kind krijgt het woord te horen en bepaalt dan uit een aantal mogelijkheden de letter die ingevuld moet worden.

deze oefeningen vervolgens zelfstandig afwerken. De resultaten (aantal fouten, snelheid en aantal keren om extra geluid gevraagd) worden automatisch geregistreerd en kunnen achteraf bekeken worden, waarna een eventuele bijstelling volgt.

De belangrijkste instructiekenmerken van COPAL zijn de volgende:

- Stimuli worden, al dan niet simultaan, via twee modaliteiten aangeboden: visueel-orthografisch (op het scherm) en auditief (per koptelefoon).
- Het programma is interactief: antwoord wordt gegeven door middel van de muis (matchingstaken) of het toetsenbord (typetaken).
- Stimulering van cognitieve activiteiten (bijvoorbeeld eerst zelf analyseren, alvorens op gelijkenis te kunnen beoordelen/matchen, over- of natypen).
- Gepresenteerd worden hele woorden of delen van woorden, alle hoogfrequent.
- Visueel-orthografische stimuli kunnen gemaskeerd (zgn. 'flitswoorden') of niet gemaskeerd worden aangeboden. In het laatste geval blijven de woorden op het scherm staan totdat het antwoord is gegeven.
- Oefeningen kunnen herhaald worden aangeboden totdat er sprake is van beheersing.
- Feedback als controle op correctheid van antwoord (auditief; visueel; beloning).
- Woordklassen zijn één- en twee-lettergrepig en variëren van MKM-woorden tot open lettergrepen.

Trainingsprocedure. Na de voortoets trainden de leerlingen gedurende een periode van 12 weken dagelijks tien minuten met het programma, gevolgd door de natoets. Het oefenpakket was gelijk wat betreft de keuze van het soort oefeningen, maar de keuze van woordklassen en moeilijkheidsniveaus was per individu verschillend, afhankelijk van zijn beheersingsniveau. Voor de vaststelling van het individueel instapniveau werd gebruik gemaakt van de resultaten op de COTALtoets (Van der Linde & Den Otter, 1991). Op basis van de door de computer bijgehouden vorderingen werden eventueel woordklassen en moeilijkheidsniveaus iedere twee á drie weken bijgesteld. De leerlingen hebben gedurende het experiment in totaal gemiddeld 293 minuten (*sd*=88) geoefend met het programma. Er waren geen verschillen in de oefentijd tussen de groepen.

Apparatuur. Het trainingsprogramma werd evenals de bovengenoemde computergestuurde toetsen aangeboden met behulp van een microcomputer en een koptelefoon (Apple Macintosh Plus met een 20 Mb harddisk).

4 Resultaten

Op iedere toets zijn variantie-analyses met een 2 (groep) × 2 (tijdstip) design uitgevoerd, waarin groep als tussen-proefpersoon variabele en tijdstip als binnen-proefpersoonvariabele. De analyses wijzen uit dat computergestuurde training effectief is. Voor COTAL (gesommeerd over de vijf deeltoetsen) was er een significant hoofdeffect voor tijdstip: $F(1,123)=173.2$; $p<.0001$. Voor de EMT was dit eveneens het geval ($F(1,123) = 212.2$; $p<.0001$), zij het dat er bovendien een significant interactie effect is tussen groep en tijdstip: de zwakke lezers gingen er meer op vooruit dan de dyslectici: $F(1,123)=26.6$; $p<.0001$. Voor AVI gold hetzelfde: een significant hoofdeffect voor tijdstip ($F(1,123)=173.1$; $p<.0001$), en een interactie effect tussen groep en tijdstip: de progressie van de zwakke lezers was groter $F(1,123)=7.0$; $p<.01$.

Is het effect van intensieve computergestuurde instructie afhankelijk van de ernst van de leesproblemen?

De eerste onderzoeksvraag kan worden beantwoord door de groepen op de natoets met elkaar te vergelijken. In de Tabellen 4a en 4b zijn de gemiddelden en standaarddeviaties op voor- en natoets alsmede de natoetsverschillen tussen beide groepen weergegeven.

Uit deze tabellen blijkt dat de COTAL-scores op de voortoets nogal verschillen tussen beide groepen. Om na te gaan of het effect van intensieve computergestuurde instructie afhankelijk is van de ernst van de leesproblemen zijn daarom op de natoetsscores van de gebruikte tests covariantie-analyses uitgevoerd, met beide groepen als factoren en de voortoetsscore als covariaat.

ZWAK was in nagenoeg alle natoetsen beter dan DYS. Significant hoger waren de scores op de EMT ($F(1,123)=28.1$; $p<.0001$) en de

Tabel 4a

Gemiddelden en standaarddeviaties (tussen haakjes) voor de EMT, AVI, COTALsom hoogfrequent+pseudo, COTALsom hoogfrequent en de afzonderlijke taken met hoogfrequente woorden; n ZWAK=48; n DYS=77. Verder natoetsverschillen tussen de groepen met de voortoetsscore als covariaat(F/p)

		ZWAK		DYS			
Variabele	Max.	Voor	Na	Voor	Na	*F*	*p*
EMT		27.9 (11.4)	40.5 (13.0)	24.2 (8.4)	30.2 (10.0)	28.1	****
AVI[1]		3.4 (1.4)	4.8 (1.9)	2.7 (1.5)	3.4 (1.8)	6.1	***
COTAL som hf+ps	70	22.4 (15.9)	34.6 (17.9)	17.4 (10.6)	27.0 (12.8)	3.2	**
		Hoogfrequente woorden					
COTAL som hf	40	13.3 (10.3)	20.9 (10.1)	10.2 (7.0)	16.1 (8.0)	4.5	**
A-V	12	5.1 (5.0)	7.5 (4.5)	5.1 (4.6)	7.4 (3.9)	0	
L	4	2.8 (1.3)	3.6 (0.8)	2.2 (1.3)	2.3 (1.3)	31.2	****
FL	8	2.5 (2.4)	4.4 (2.8)	0.9 (1.4)	2.4 (2.1)	2.5	*
V-T	8	1.2 (2.0)	2.7 (2.4)	0.6 (0.9)	1.5 (1.8)	4.8	**
A-T	8	1.7 (2.2)	2.8 (2.7)	1.4 (1.9)	2.4 (2.1)	0.5	

* $p<.10$
** $p<.05$
*** $p<.01$
**** $p<.001$
1. *n ZWAK*=33
n DYS=63

Tabel 4b

Gemiddelden en standaarddeviaties (tussen haakjes) voor de COTALsom pseudo en de afzonderlijke taken met pseudo-woorden; n ZWAK=48; n DYS=77. Verder natoetsverschillen tussen de groepen met de voortoetsscore als covariaat(F/p)

		Pseudo-woorden					
		ZWAK		DYS			
Variabele	Max.	Voor	Na	Voor	Na	*F*	*p*
COTAL som ps	40	9.1 (6.9)	13.8 (8.7)	7.2 (4.8)	10.9 (6.0)	1.6	
A-V	12	4.3 (3.7)	5.5 (3.7)	4.6 (3.5)	6.0 (3.3)	0.5	
L	4	2.1 (1.2)	2.5 (1.0)	1.5 (1.2)	1.9 (1.2)	5.0	**
FL	8	1.4 (1.6)	3.0 (2.4)	0.5 (0.8)	1.2 (1.3)	9.9	***
V-T	8	0.7 (1.5)	1.5 (2.1)	0.3 (0.7)	0.7 (1.1)	3.5	**
A-T	8	0.5 (1.3)	1.2 (1.9)	0.4 (1.0)	1.2 (1.8)	0	

** $p<.05$
*** $p<.01$

AVI ($F(1,123)=6.1$; $p<.01$; beide eenzijdig getoetst). Computergestuurde oefening op woordniveau leidde dus tot verschil in generalisatie naar het algemeen woordleesniveau en naar het zinsniveau. Vertaald naar de EMT-normscores betekent dit dat de zwakke lezers er na zes maanden training op woordniveau (inclusief voor- en natoets) evenveel maanden op vooruit zijn gegaan in hun algemene woordleesniveau als bij een toename van de leeftijd met zes maanden mag worden verwacht bij normale lezers (van EMT 28 naar EMT 41). De dyslectische lezers gingen er in dezelfde periode slechts ongeveer drie maanden op vooruit. De generalisatie naar het zinsniveau (AVI) verschilde eveneens significant tussen beide groepen, maar opvallend was dat beide groepen minder vooruitgingen dan men zou mogen verwachten bij een normale leesontwikkeling. De zwakke lezers gingen er in deze periode ongeveer vier maanden op vooruit, de dyslectici ongeveer drie maanden.

Op de gesommeerde natoetsscore (COTALsom hf+ps) van alle hoogfrequente en pseudo-woordtaken (met de gesommeerde voortoetsscore als covariaat), scoorde ZWAK significant hoger dan DYS ($F(1,23)=3.2$; $p<.05$); het verschil op de gesommeerde hoogfrequente score (COTALsom hf) was eveneens significant ($F(1,123)=4.5$; $p<.05$). Op taakniveau bleek ZWAK beter te zijn in het benoemen van hoogfrequente woorden, L ($F(2,46)=31.2$; $p<.0001$) en pseudo-woorden ($F(1,123)=5.0$; $p<.05$); in het benoemen van geflitste pseudo-woorden ($F(1,123)=9.9$; $p<.001$); en in het natypen van geflitste hoogfrequente woorden, V-T ($F(1,123)=4.8$; $p<.01$) en pseudo-woorden ($F(1,123)=3.5$; $p<.05$). Ten slotte was er nog een tendens in dezelfde richting voor het benoemen van hoogfrequente flitswoorden, FL ($F(1,123)=2.5$; $p<.06$).

Samengevat kunnen we dus stellen dat zwakke lezers blijkbaar meer profiteren van computergestuurd oefenen op woordniveau dan dyslectici. Het effect wordt in nagenoeg alle taken gevonden zowel met hoogfrequente als pseudo-woorden. Alleen het auditief-visueel matchen (A-V) en het natypen van auditief aangeboden woorden (A-T) laat geen differentiële effecten zien.

Is er een effect van intensieve computergestuurde instructie op de ontwikkeling van de lexicale en fonologische verwerking van woorden?

Zoals reeds in paragraaf twee gesteld behelst deze vraag of training leidt tot betere, dat wil zeggen accuratere en snellere lexicale en fonologische verwerking. Het effect op de accuratesse van beide verwerkingswijzen is nagegaan door bij beide groepen het percentage leerlingen dat tot beheersing (het wel of niet halen van het 90% criterium) van de respectievelijke woordklassen in de benoemtaak met onbeperkte aanbieding (L) komt, op voor- en natoets te vergelijken. Het effect op snelheid is gemeten door bij beide groepen het percentage leerlingen dat gemiddeld over de woordklassen tot beheersing komt van de benoemtaak met onbeperkte aanbieding (L), in voor- en natoets te vergelijken met het percentage dat op de benoemtaak met flitsaanbieding (FL) tot beheersing komt: het effect van presentatietijdverkorting. Daarnaast kan het effect worden gedifferentieerd naar de toenemende fonologische en orthografische complexiteit van de verschillende woordklassen (zie par. 2).

Accuratesse. Figuur 1 toont dat computergestuurde training wat betreft accuratesse maar bij weinig dyslectici tot een verbetering van de lexicale woordverwerking heeft geleid. Bij de twee-lettergrepige woorden was er zelfs sprake van een lichte daling van het aantal leerlingen dat tot beheersing komt (van 49% naar 46%). Toetsing met de McNemar test voor significantie van veranderingen (Siegel & Castellan, 1988) wees uit dat er in geen enkele woordklasse een significante toename van het aantal leerlingen was tussen voor- en natoets. Zou dit voor de MKM-woorden nog verklaard kunnen worden door plafond-effecten, voor de andere klassen was dit zeker niet het geval. Ondanks intensieve oefening bleek slechts 22% van de leerlingen in staat woorden met open lettergrepen, die onbeperkt worden aangeboden, correct te kunnen benoemen.

Bij de zwakke lezers was de toename van het aantal leerlingen daarentegen wel significant bij de MKMM/MMKM-woorden (Mc Nemars $\chi^2(df\ 1)=6.13$; $p<.02$), de twee-lettergrepige woorden ($\chi^2(df\ 1)=6.75$; $p<.01$) en de woorden met open lettergrepen ($\chi^2(df\ 1)=11.5$; $p<.001$). Dat er bij de MKM-woorden

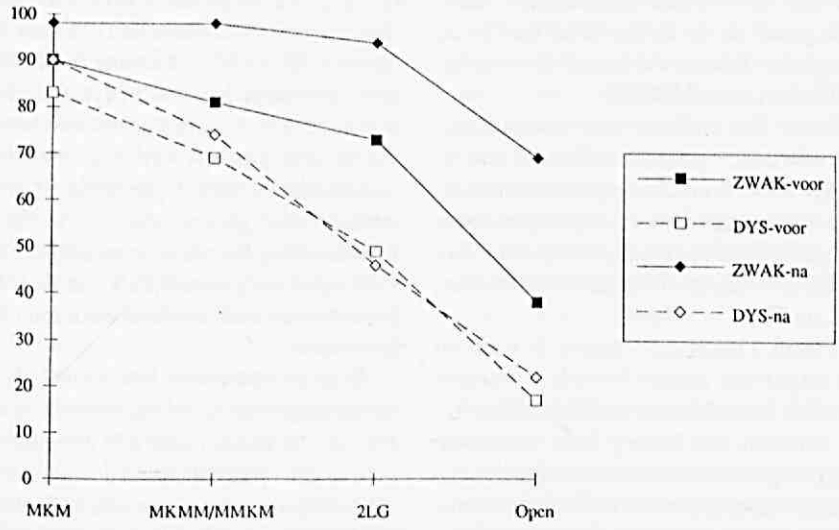
geen significante toename was, kan waarschijnlijk toegeschreven worden aan plafondeffecten: 90% van de leerlingen beheerste deze categorie reeds op de voortoets. Na de training waren er nog maar enkele zwakke lezers die de eerste drie woordklassen niet beheersten, terwijl dit voor slechts 31% met betrekking tot de open lettergrepen gold.

Wat betreft de ontwikkeling van de lexicale woordverwerking bleek de instructie niet te differentiëren naar fonologische en orthografische complexiteit van de klassen. Het verval op beide factoren bleef, afgezien van de waarschijnlijk door plafond-effecten veroorzaakte niet significante vooruitgang in de MKM-woordklasse, voor en na de training ongeveer gelijk bij beide groepen. Afgezien van de instructie bleek het verschil tussen MKM en MKMM/MMKM op voor- en natoets bij de dyslectische kinderen groter dan bij de zwakke lezers, wat een ondersteuning geeft aan de hypothese dat dyslectici meer problemen hebben met een toenemende fonologische complexiteit dan zwakke lezers (Van der Leij, in press). Er was echter geen verschil tussen tweelettergrepige woorden en woorden met een open lettergreep, afgezien van het feit dat dyslectische leerlingen meer moeite hadden met beide woordklassen.

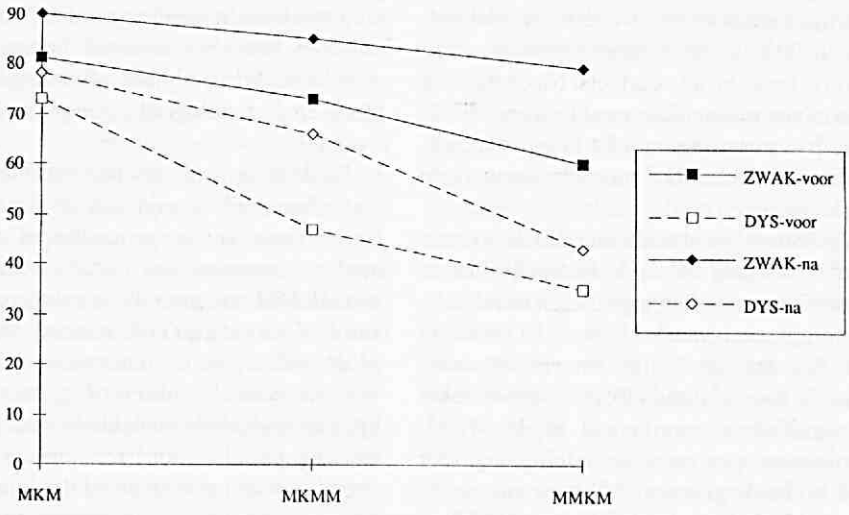
In Figuur 2 zijn de effecten van de training op de ontwikkeling van een accuratere fonologische woordverwerking weergegeven. Training met hoogfrequente woorden lijkt te generaliseren naar het accuraat benoemen van pseudo-woorden, althans afhankelijk van de groep en de fonologische complexiteit van de woorden.

Bij de groep dyslectische lezers bleek het effect afhankelijk te zijn van de fonologische complexiteit en te generaliseren naar het hardop benoemen van pseudo-woorden met een MKMM-structuur (beheersingspercentage van 47% vooraf naar 66% achteraf; McNemars $\chi^2(df\,1)=6.22$; $p<.02$), maar niet naar MMKM-woorden (van 35% naar 43%; χ^2=n.s.). Mogelijk was de dubbele medeklinker aan het begin van een pseudo-woord fonologisch (nog) te complex voor het merendeel der kinderen om tot accurate benoeming te komen. In geen van beide groepen viel er een significante toename te noteren van het aantal leerlingen dat tot beheersing kwam in de MKM-categorie. De verklaring hiervoor moet, gezien de relatief hoge percentages, gezocht worden in plafondeffecten en/of te grote fonologische problemen, althans voor een aantal kinderen en mogelijk voor meer kinderen in de dyslectische groep.

De zwakke lezers boekten in beide andere woordklassen wel vooruitgang. Was dit in de MKMM-woordklasse niet meer dan een tendens ($\chi^2(df\,1)=3.13$; $p<.10$), bij de MMKM-



Figuur 1. Percentages leerlingen per groep die de verschillende woordklassen beheersen in de benoemtaak met onbeperkte aanbieding (L; hoogfrequente woorden) in voor- en natoets



Figuur 2. Percentages leerlingen per groep die de verschillende woordklassen beheersen in de benoemtaak met onbeperkte aanbieding (L; pseudo-woorden) in voor- en natoets

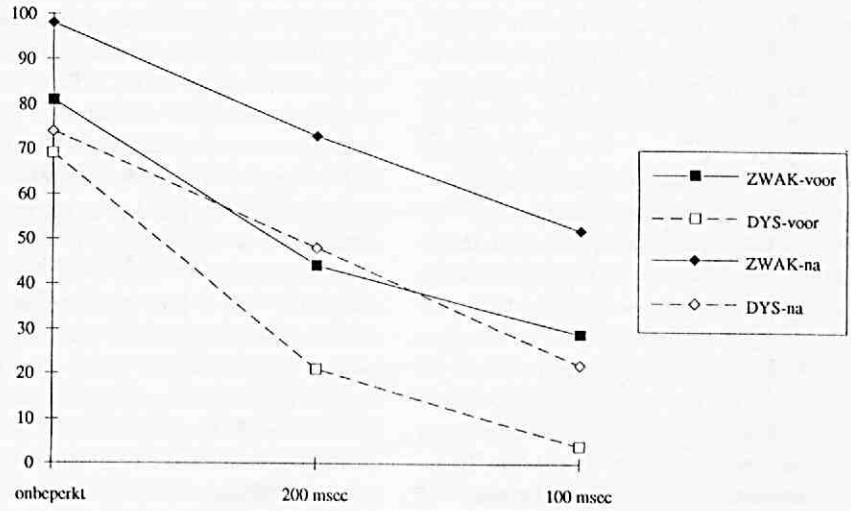
woorden was de progressie wel significant ($\chi^2(df\,1)=6.67$; $p<.01$). Hoewel de vooruitgang het grootst was bij de MMKM-woorden (van 60% naar 79%), werd niet duidelijk of dit samenhangt met de fonologische complexiteit van de gebruikte klassen. De verschillen tussen de woordklassen zijn na de training wel kleiner geworden, de meeste leerlingen hebben geleerd deze een-lettergrepige woorden fonologisch accuraat te verwerken.

Op de natoets bleek volgens verwachting de invloed van toenemende fonologische complexiteit groter in de dyslectische dan in de zwakke groep, althans wat betreft de vergelijking MKMM versus MMKM.

Snelheid. De invloed van computergestuurde instructie op de ontwikkeling van de snelheid in lexicale en fonologische woordverwerking is nagegaan door beide groepen lezers in voor- en natoets te vergelijken op het effect van presentatietijdverkorting in beide benoemtaken (L en FL).

Uit Figuur 3 blijkt dat wanneer de voor- en natoets vergeleken worden in beide groepen en onder iedere conditie meer leerlingen hoogfrequente woorden snel hebben leren verwerken en ondanks presentatieduurverkorting tot een correcte benoeming komen. Deze toename was slechts onder één conditie niet significant (McNemars χ^2) en wel bij de onbeperkte aanbieding van woorden in de groep dyslectici (L; van 69% naar 74% leerlingen). Zoals reeds geconstateerd werd in deze groep op de natoets geen enkele woordklasse door significant meer lezers beheerst in vergelijking met de voortoets. Na de training bleken er (ook relatief ten opzichte van de onbeperkte conditie) significant meer dyslectici te zijn die ondanks verkorting van de presentatieduur, tot een accurate benoeming van de woorden komen. In de 200 msec. conditie steeg het percentage leerlingen dat tot beheersing kwam van 21% naar 48% ($\chi^2(df\,1)=17.34$; $p<.001$); in de 100 msec. conditie ging dit percentage van 4% naar 22% (χ^2 ($df\,1)=8.69$; $p<.01$). Oefening in het snel lexicaal verwerken, had dus bij dyslectische lezers een direct effect op een accuratere benoeming van hoogfrequente woorden die maar heel kort werden aangeboden. Opgemerkt zij echter dat, ondanks deze grotere proportie leerlingen die tot beheersing kwam, in de moeilijkste conditie (100 msec.) nog steeds 78% van de leerlingen grote moeite had om deze woorden correct te benoemen.

In de groep zwakke lezers bleek de proportie leerlingen die na het experiment tot accurate benoeming kwam onder alle condities significant te zijn toegenomen. In de onbeperkte conditie steeg het percentage van 81% naar 98 (χ^2 ($df\,1)=6.13$; $p<.02$), bij verkorting van de stimuluspresentatie tot 200 msec. van 44% naar 73% ($\chi^2(df\,1)=8.69$; $p<.01$), de 100 msec. con-



Figuur 3. Percentages leerlingen per groep (gesommeerd over woordklassen) die de verschillende aanbiedingstijden beheersen in benoemtaken (L, FL; hoogfrequente woorden) in voor- en natoets

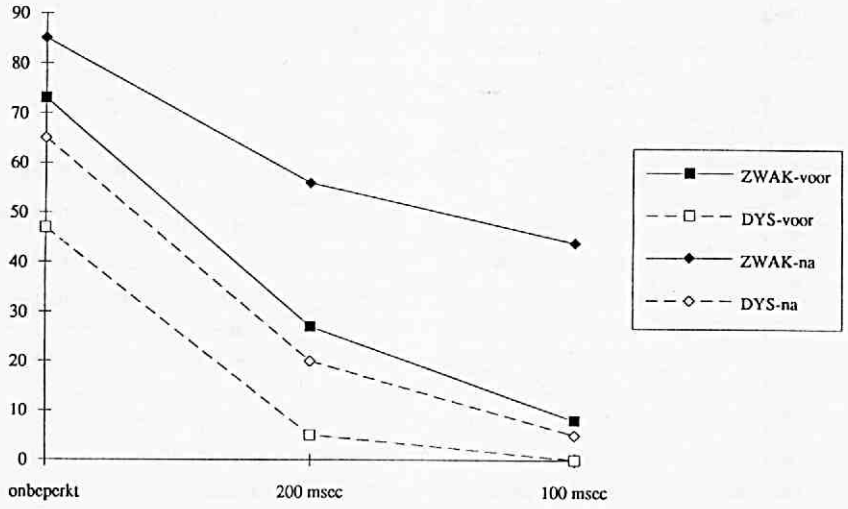
ditie tenslotte werd op de natoets door 52% van de leerlingen beheerst in vergelijking met 29% op de voortoets ($\chi^2(df\ 1)=6.67$; $p<.01$). Meer dan de helft van de groep zwakke lezers bleek dus na de training zelfs onder de moeilijkste conditie de hoogfrequente woorden nog correct te kunnen benoemen, in tegenstelling tot de slechts 22% bij de dyslectici.

Figuur 4 toont dat oefening in het snel lexicaal verwerken van woorden ook generaliseert naar een snellere fonologische woordverwerking, in casu het hardop benoemen van pseudowoorden. Deze effecten zijn echter niet even groot in beide groepen en niet onder alle condities aanwezig. Het aantal dyslectische leerlingen dat tot beheersing kwam, werd zowel in de onbeperkte ($\chi^2(df\ 1)=4.63$; $p<.05$; van 47% naar 65%) als in de 200 msec. conditie ($\chi^2(df\ 1)=7.69$; $p<.01$: van 5% naar 20%) significant groter, echter niet in de moeilijkste conditie (van 0% naar 5%; χ^2 n.s.). De fonologische verwerking van pseudo-woorden bleef bij verreweg de meeste dyslectici dus ondanks intensieve oefening met bestaande woorden een groot struikelblok zodra ze onder tijdsdruk verwerkt moesten worden.

Bij de zwak lezende kinderen waren er daarentegen wel grote generalisatie-effecten hetgeen bleek uit de toename van het aantal lezers dat de moeilijkere condities beheerste, zowel absoluut als relatief ten opzichte van de onbeperkte aanbieding. Was de toename in de onbeperkte conditie niet meer dan een tendens ($\chi^2(df\ 1)=3.13$; $p<.10$; van 73% vooraf naar 85% achteraf; mogelijk als gevolg van plafondeffecten en/of enkele hardleerse proefpersonen), in de beide geflitste aanbiedingen nam het significante waarden aan. In de 200 msec. conditie las 56% van de leerlingen op de natoets de pseudo-woorden correct in vergelijking met slechts 27% op de voortoets ($\chi^2(df\ 1)=8.69$; $p<.01$). In de moeilijkste conditie is dit verschil nog groter: slechts 8% vooraf tegen 44% achteraf ($\chi^2(df\ 1)=15.06$; $p<.001$). Hoewel verreweg de meeste zwak lezende kinderen deze bemoeilijkende condities op de voortoets niet aankonden, is de toename van dit aantal, ook in vergelijking met de onbeperkte aanbieding groot. De conclusie dat veel van deze kinderen de snelheid waarmee ze pseudowoorden verwerken in de tussenliggende periode behoorlijk hebben ontwikkeld, lijkt dan ook gerechtvaardigd.

5 Discussie en conclusies

In het algemeen kan worden geconcludeerd dat de resultaten van deze voorstudie positief zijn. Interactieve computergestuurde oefening blijkt inderdaad uitvoerbaar en leidt - zo is gebleken uit het commentaar van de leerkrachten - niet



Figuur 4. Percentages leerlingen per groep (gesommeerd over woordklassen) die de verschillende aanbiedingstijden beheersen in benoemtaken (L, FL; hoogfrequente woorden) in voor- en natoets

tot verveling bij de leerlingen of grote problemen bij de toepassing. Daarnaast zijn de effecten die deze eerste beproeving van het programma te zien gaf, bemoedigend. Zowel de dyslectische als de zwakke lezers lijken te profiteren van deze vorm van computergestuurde training. Beide groepen lezers gingen in nagenoeg alle toetsen significant vooruit tussen voor- en natoets. Deze vooruitgang betekent dat er aanwijzingen zijn dat de drie vormen van generalisatie-effecten waarin we geïnteresseerd waren, inderdaad optraden. De eerste vorm bleek uit de vooruitgang die beide groepen boekten in de verwerking van de woordklassen die in de voor- en natoets op soortgelijke computergestuurde wijze als in de oefening werden aangeboden. Bovendien boekten beide groepen een significante vooruitgang in de verwerking van niet bestaande woorden, terwijl de training slechts bestaande woorden betrof. De betere prestaties op niet-geoefende taken (de vooruitgang op de EMT en de AVI) wijzen er op dat ook de tweede vorm van generalisatie optrad: die naar taken waarbij de computer geen rol speelde en bovendien de woorden in een andere vorm (rijen resp. verhaaltjes) werden aangeboden. Tot slot is er de derde vorm van generalisatie: computergestuurde training van het stillezen lijkt een remediërend effect te hebben op het hardop lezen.

Op deze plek dient echter de kanttekening te worden geplaatst dat op basis van deze bevindingen niet kan worden bepaald hoe robuust deze generalisatie-effecten zijn. In dit design zijn namelijk geen vergelijkingen getrokken met groepen lezers die niet op een dergelijke wijze getraind hebben (zgn. controlegroepen). Daardoor is het niet uit te maken in hoeverre de vooruitgang toe te schrijven is aan 'natuurlijke' ontwikkeling (dat wil zeggen ontwikkeling zonder intensieve oefening). Verder onderzoek waarin wel controlegroepen zijn opgenomen zal moeten uitwijzen in hoeverre de generalisatie-effecten inderdaad toe te schrijven zijn aan de computergestuurde oefeningen.

Het design kan echter wel een antwoord geven op de vraag of de effectiviteit samenhangt met de ernst van de leesproblemen. Dat is inderdaad het geval. Zwakke lezers met ruim een jaar achterstand op het voor hun leeftijd verwachte algemene leesniveau, blijken meer te profiteren van de training dan dyslectische leerlingen met meer dan twee jaar achterstand. Het feit dat de zwakke lezers een grotere vooruitgang boekten op beide flitstaken (FL en V-T) dan de dyslectici, duidt erop dat zwakke lezers meer vaardigheid hebben ontwikkeld in het snel waarnemen en verwerken van woorden. Aangezien juist het steeds sneller verwerken een kenmerk van normale leesontwikkeling is (Van der Leij, in press), kunnen we stellen dat de door de training verkregen rela-

tief grotere progressie bij de zwakke lezers tendeert in de richting van een normale ontwikkeling. Dit wordt eveneens ondersteund door het feit dat de zwakke lezers er na zes maanden training op woordniveau evenzoveel maanden op vooruitgaan als bij een toename van de leeftijd met zes maanden mag worden verwacht bij normale lezers (zoals blijkt uit de toename op de EMT). Indien de hypothese van het toenemende verschil in leesontwikkeling tussen normale en zwakke lezers juist is (het 'Mattheuseffect'; Stanovich, 1986), dan zouden zwakke lezers, zonder extra instructie geen gelijke maar een tragere progressie moeten vertonen. Op zichzelf is de vooruitgang van de zwakke lezers door toedoen van intensieve instructie, op te vatten als een positief resultaat. De eerder beschreven hypothese dat de oorzaak van zwak lezen beter te kenmerken is door een achterstand in leesontwikkeling dan door een specifiek, structureel tekort, krijgt door dit resultaat ook steun: wanneer deze kinderen extra instructie krijgen, halen ze hun achterstand in.

Wat betreft de dyslectische lezers is de situatie minder rooskleurig. Weliswaar heeft intensief, computergestuurd oefenen een positief effect, maar de vooruitgang is minder groot dan bij de zwakke lezers. Daarbij duidt vooral de minder grote vooruitgang in het verwerken van flitswoorden (zowel hoogfrequente als pseudowoorden) op beperkingen in de snelle verwerking van visueel-orthografische stimuli bij dyslectische kinderen. Uit ander onderzoek blijkt echter dat het snel analyseren en verwerken te stimuleren valt bij zeer zwakke lezers (Frederiksen et al, 1985a; Van den Bosch, Van Bon & Schreuder, 1990; Klapwijk & Das-Smaal, 1990; Van den Bosch, 1991; Yap & Van der Leij, in voorbereiding). In tegenstelling tot COPAL was de instructie in deze onderzoeken uitsluitend gericht op snelle verwerking (flitstaken). Wellicht is een dergelijke concentratie van de instructie op specifieke deelvaardigheden effectiever dan meer gevarieerde oefening zoals bij COPAL het geval is. Het onderzoek van Yap en Van der Leij (in voorbereiding) waarin flitsoefeningen vergeleken werden met COPAL, duidt er echter op dat het grotere effect van deze specifieke instructie slechts van korte duur is en na verloop van tijd verdwijnt. Vooralsnog houden wij het er op dat de resultaten van het trainingsexperiment waarover in dit artikel gerapporteerd wordt, het bestaan van een structureel tekort in informatieverwerken van dyslectische kinderen niet tegenspreken, terwijl in het geval van de zwakke lezers de achterstandshypothese duidelijke steun krijgt.

De vraag of de ontwikkeling van de lexicale en fonologische verwerking gestimuleerd wordt, lijkt voor beide groepen lezers bevestigend beantwoord te kunnen worden. Op de eerste plaats is er een *accuratere* verwerking van woorden en pseudo-woorden. Vermeld zij echter dat het effect groter is bij zwakke lezers. In nagenoeg alle gebruikte woordklassen nam het aantal zwakke lezers dat tot beheersing kwam significant toe. Verreweg de meesten bleken aan het eind van het experiment in staat woorden en pseudo-woorden op een accurate wijze te verwerken. Bij de dyslectici daarentegen was deze toename alleen significant bij pseudo-woorden met een MKMM-structuur.

Wat betreft de invloed van computergestuurde instructie op de ontwikkeling van de *snelheid* van lexicale en fonologische woordverwerking, kunnen we stellen dat beide groepen erop vooruit zijn gegaan aan het eind van het experiment. Analoog aan de accuratessekwestie is ook hier de progressie bij de zwakke lezers groter dan bij de dyslectici. De grote toename van het aantal leerlingen dat bij de pseudo-woorden tot een accurate respons komt ondanks de tijdsdruk (100 msec.), biedt perspectief voor deze zwakke lezers. Het steeds sneller verwerken is immers een belangrijk kenmerk van normale leesontwikkeling (Van der Leij, in press). Het aantal dyslectische lezers dat na de training tot beheersing van de verschillende condities komt, neemt weliswaar toe, maar blijft in de 100 msec./pseudowoorden conditie ver achter bij de zwak lezende groep, zodat betwijfeld mag worden dat ze deze conditie zullen leren beheersen.

Computergestuurde instructie heeft bij dyslectische lezers wel enig effect op de ontwikkeling van de lexicale en fonologische verwerking van woorden, maar dit effect wordt beperkt door de problemen die deze leerlingen ondervinden van enerzijds toenemende fonologische complexiteit van de woorden, anderzijds van presentatietijdverkorting, in casu de toenemende tijdsdruk waaronder de woorden moeten worden verwerkt. Als geautomatiseerd

lezen gedefinieerd wordt als het snel en accuraat verwerken van woorden en pseudowoorden ondanks bemoeilijkende condities, dan mag worden betwijfeld of deze fase, gegeven de gevonden resultaten, bereikt zal worden door de dyslectische lezers, met uitzondering wellicht voor wat betreft de eenvoudigste woorden. Hun problemen hebben zowel betrekking op de kwaliteit van de fonologische code waarin woorden voor verdere verwerking dienen te worden omgezet (af te lezen uit de negatieve invloed van toenemende fonologische complexiteit van pseudo-woorden) als op de snelheid van verwerking (zoals blijkt uit de negatieve invloed van presentatietijdverkorting), en derhalve op beide pijlers onder automatisering (verg. ook Yap (1993) en Van der Leij (in press)).

Deze resultaten tonen overigens geenszins de superioriteit van computergestuurde instructie en oefening ten opzichte van andere vormen van individuele instructie aan. Omdat in deze studie geen vergelijkingen zijn gemaakt met andere vormen van intensieve training, zoals remedial teaching, kunnen we geen antwoord geven op de vraag welke vorm het grootste effect sorteert. Toch duiden de hier gevonden resultaten erop dat interactieve, remediërende computergestuurde instructie haar vruchten afwerpt en als zodanig een hulpmiddel kan zijn in het leesonderwijs aan leerlingen met leesproblemen. Ook praktisch gesproken kan dit resultaat positief worden genoemd: computergestuurde instructie is immers organisatorisch en financieel een aantrekkelijk alternatief voor remedial teaching.

Concluderend kan gesteld worden dat het gebruik van COPAL zinvol lijkt te zijn bij zwakke lezers die nog geen al te grote achterstand hebben. Wat betreft de dyslectische lezers (althans, de groep van dyslectische lezers die dusdanig complexe leerproblemen heeft dat zij aangewezen is op het LOM-onderwijs) is het echter heel aannemelijk dat hun mogelijkheid om de technische leesvaardigheid volledig te automatiseren, structureel beperkt is. Wellicht zou het in hun geval gewenst zijn om remediërende oefeningen, zoals vormgegeven in COPAL, aan te vullen met compenserende oefeningen die veeleer strategieën betreffen om de betekenis van de tekst te achterhalen. Ook dat kan in computergestuurde vorm (verg. bijv. Walraven & Reitsma, 1991). Uiteraard is daarmee nog niet gezegd dat de combinatie van remediërende en compenserende oefeningen dyslectische leerlingen inderdaad tot een zodanige leesvaardigheid kan brengen dat de ontwikkeling van de functionele geletterdheid zonder al te veel problemen kan plaats vinden (verg. Van der Leij, 1992).

Allereerst is het echter zaak om de effectiviteit van COPAL in een grootschaliger studie opnieuw te beproeven. Daarbij zal ook aandacht besteed worden aan de vergelijking met controlegroepen die geen training krijgen en andere experimentele condities dan de twee die in het onderhavige experiment zijn gebruikt. Te zijner tijd hopen we daarover te rapporteren.

Literatuur

Berg, R. M. van den, & Lintelo, H. G. te (1977). *A.V.I.-pakket*. Den Bosch: K.P.C..

Boland, T. (1991). *Lezen op termijn* (dissertatie). Nijmegen: Katholieke Universiteit.

Bosch, K. van den, Bon, W. H. J. van, & Schreuder, R. (1990). Flash cards; een vergelijking van toepassingsprocedures. *Tijdschrift voor Orthopedagogiek, 29,* 233-244.

Bosch, K. van den (1991). *Poor reader's decoding skills.* (dissertation.). Nijmegen: Katholieke Universiteit.

Bouma, H., & Legein, C. P. (1980). Dyslexia: a specific recoding deficiency. An analysis of response latencies for letters and words in dyslectics and average readers. *Neuropsychologica, 18,* 285-298.

Brus, B. T., & Voeten, M. J. M. (1973). *Een-minuuttest, Vorm A en B.* Nijmegen: Berkhout.

Daal, V. H. P. van, Leij, A. van der, Bakker, N. C. M., & Reitsma, P. (1987). Een Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen. *Pedagogische Studiën, 64,* 364-376.

Daal, V. H. P. van, Leij, A. van der, & Geervliet-van der Hart, J. A. (1989). Het effect van type-oefeningen op het leren lezen en spellen van kinderen met ernstige leesproblemen. *Pedagogische Studiën, 66,* 185-192.

Dumont, J. J. (1990). *Dyslexie. Theorie, diagnostiek, behandeling.* Rotterdam: Lemniscaat.

Frederiksen J. R., Warren, B. M., & Roseberry, A. S. (1985a). A componential approach to training reading skills: Part 1. Perceptual units training. *Cognition and Instruction, 2,* 91-130.

Frederiksen J.R., Warren, B.M., & Roseberry, A.S. (1985b). A componential approach to training reading skills: Part 2. Decoding and use of context. *Cognition and Instruction, 2,* 271-338.

Klapwijk, M.J.G., & Das-Smaal, E. (1990). *Leesvaardigheidstraining volgens een componentenbenadering.* Amsterdam: Vrije Universiteit.

LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Leij, A. van der (1983). *Ernstige leesproblemen. Een onderzoek naar mogelijkheden tot differentiatie en behandeling* (dissertatie). Lisse: Swets & Zeitlinger.

Leij, A. van der (1990). De constructie van een computergestuurd programma voor aanvankelijk lezen. *Tijdschrift voor Orthopedagogiek, 29,* 245-257.

Leij, A. van der (1992). Risico op functionele ongeletterdheid in de basisschool: een taxonomie van oorzaken en een beschrijving van mogelijkheden tot aangepaste interventie. *Pedagogische Studiën, 69,* 352-370.

Leij, A. van der. The development of the autonomous lexicon of reading disabled students. Paper presented at the Rodin Conference Bern. (in press; to be published in the proceedings).

Leij, A. van der, & Reitsma, P. (1990). The development of a computer-assisted training program for reading-disabled children. In L. de Leeuw, R.J. Simons & J.M. Pieters (Eds.) *Computers and Instructional Research* (pp. 115-127). Lisse: Swets & Zeitlinger.

Leij, A. van der, Smeets, H., & Daal, V.H.P. van (1990). De ontwikkeling van de Computergestuurde Toets voor Aanvankelijk Lezen. *Pedagogische Studiën, 67,* 163-178.

Lieshout, E.C.D.M. van, (1990). Computerondersteund onderwijs voor kinderen met ernstige leerproblemen. *Tijdschrift voor Orthopedagogiek, 29,* 223-232.

Linde-Kaan, A. van der, & Otter, M. den (1991). *Effecten van computergestuurde instructie bij kinderen met leesproblemen.* (ongepubliceerd manuscript) Amsterdam: Vrije Universiteit.

Olson, R.K., Kliegl, R., Davidson, B.J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In T.G. Waller (Ed.) *Reading research: advances in theory and practice, Vol. 4,* (pp. 1-64). New York: Academic Press.

Perfetti, C.A. (1985). *Reading ability.* New York: Oxford University Press.

Rack, J.P., Snowling, M., & Olson, R.K. (1992). The nonword reading deficit in developmental dyslexia: a review. *Reading Research Quarterly, 27*(1), 28-53.

Samuels, S.J. (1985). Automaticity and repeated reading. In J. Osborn, P.T. Wilson & R.C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 215-223). Lexington: Lexington Books.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: Detection, search and attention. *Psychological Review, 84,* 1-66.

Siegel, S., & Castellan, N.J., JR. (1988). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Snowling, M. (1987). *Dyslexia: A cognitive developmental perspective.* Oxford: Basil Blackwell Ltd..

Spear, L.C., & Sternberg, R.J. (1987). An information-processing framework for understanding reading disability. In S.J. Ceci (Ed.) *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (pp. 3-31). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407.

Stanovich, K.E., Nathan, R.G., & Zolman, J.E. (1988). The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. *Child Development, 59,* 71-86.

Treiman, R., & Hirsh-Pasek, K. (1985). Are there qualitative differences in reading behavior between dyslexics and normal readers? *Memory and Cognition, 13,* 357-364.

Walraven, M., & Reitsma, P. (1991). Computergestuurde cloze-oefening en begrijpend lezen. In P. Reitsma & M. Walraven (Red.), *Instructie in begrijpend lezen* (pp. 61-75). Delft: Eburon.

Yap, R.L. (1993). *Automatic word processing deficits in dyslexia: qualitative differences and specific remediation.* Dissertatie Vrije Universiteit, Amsterdam.

Yap, R.L., & Leij, A. van der. Word processing in dyslexics: an automatic decoding deficit? (in press; *Journal of reading and writing*).

Yap, R.L., & Leij, A. van der. Computergestuurde remediëring van dyslexie door het opvoeren van de herkenningssnelheid van subwoordeenheden. (in voorbereiding; aangeboden aan Pedagogische Studiën).

Auteurs

H. Smeets (1956) studeerde orthopedagogiek aan de R.U. te Leiden en werkt sinds 1988 als wetenschappelijk onderzoeker aan de Vrije Universiteit. Sinds 1989 is hij verbonden aan het NWO-project: 'Directe woordherkenning bij kinderen met leesproblemen'.

A. van der Leij (1946) is hoogleraar aan de Vrije Universiteit.

Correspondentie-adres: Vrije Universiteit, Vakgroep Pedagogiek, Van der Boechorststraat 1, 1081 BT Amsterdam.

Manuscript aanvaard 11-1-1993

Abstract

Differential effects of computer-assisted instruction on the reading ability of poor and dyslexic readers: a pilot study

H. Smeets & A. van der Leij. Pedagogische Studiën, 1993, 70, 56-72.

A pilot study is described on the effects of computer-assisted instruction on two aspects of reading ability of 48 poor and 77 dyslexic readers, eight and almost ten years of age respectively. The first research question was whether the effect was dependent on degree of reading backwardness (poor versus dyslexic). The second question concerned the effect on lexical and phonological processing (processing of high-frequency and pseudowords respectively) and on automatization of these processes. The results indicate that computer-assisted instruction was effective in both groups, but that the poor readers benefited more from the instruction. The development of (automatized) lexical and phonological processing accelerated in poor readers so that this processing showed signs of normal development. However, the dyslexic readers progressed at a slower rate. It is therefore doubtful that they will reach the stage of automatization. At the end of the article the results are discussed.

Boekbespreking

G. J. Vreeke
Zorg en rechtvaardigheid
Analyse van de Kohlberg-Gilligandiscussie
Dekker & Van de Vegt, Assen 1992 VIII + 196 pagina's
ƒ 37,50 ISBN 90 255 0076 5
(dissertatie Vrije Universiteit Amsterdam, promotor B. Spiecker, copromotor J.W. Steutel)

Kohlbergs theorie is lange tijd de dominante theorie op het gebied van de morele ontwikkeling geweest. Het is ook een produktieve theorie: ze heeft tot veel onderzoek geïnspireerd en veel kritiek en discussie uitgelokt. Een van de belangrijkste critici is Gilligan. Haar verwijt is dat Kohlbergs theorie niet de morele ontwikkeling maar slechts één aspect daarvan thematiseert, namelijk het oordelen over de rechtvaardigheid van handelingen. In haar boek *In a Different Voice* verdedigt zij de stelling dat morele problemen door mannen en vrouwen op fundamenteel verschillende wijze benaderd worden. Het denken in termen van rechtvaardigheid beschouwt zij als typisch mannelijk. Het vrouwelijk perspectief wordt veel meer gekenmerkt door persoonlijke betrokkenheid en zorg.

De vraag die in deze dissertatie centraal staat, betreft de verhouding tussen deze twee visies op moraliteit. Gaat het, zoals Gilligan zegt, om twee essentieel verschillende vormen van moreel denken? Als dat zo is, zijn beide even geschikt? Hoe moet zo'n afweging gemaakt worden? Of kan, met Kohlberg, verdedigd worden dat het rechtvaardigheidsperspectief de zorgbenadering omvat en uiteindelijk de enig acceptabele grondslag biedt? Deze vragen betreffen niet alleen de problematiek van morele ontwikkelingstheorieën, het zijn centrale ethische kwesties. Op de achtergrond spelen nog verder reikende discussies mee, die overigens door de auteur niet besproken worden, zoals bijvoorbeeld binnen het feminisme over het al dan niet accentueren van de specifiek vrouwelijke identiteit en binnen de wijsbegeerte over de reikwijdte van de idealen van de Verlichting.

In het eerste deel van zijn dissertatie geeft Vreeke een gedetailleerde uiteenzetting van de beide theorieën, in hun diverse gestalten. Hij beschrijft Gilligans kritiek dat Kohlbergs theorie het specifieke van de vrouwelijke moraliteit systematisch onderwaardeert vanuit het typisch mannelijke accent op abstracte principes en redeneringen. Vervolgens worden Kohlbergs nogal wisselende reacties op deze kritiek weergegeven.

In het tweede deel maakt de auteur duidelijk dat systematische verschillen in beoordeling van morele kwesties door mannen en vrouwen althans uit het empirisch onderzoek in het kader van Kohlbergs theorie niet naar voren komen. Wat dat betreft is het verwijt van een seksistisch vooroordeel dus ongegrond. Dat betekent natuurlijk niet dat de benadering van morele problemen niet toch heel verschillend kan zijn. Criteria worden vervolgens geformuleerd en beredeneerd om dat uit te kunnen maken.

Een zorgvuldige afweging op basis daarvan leidt in het derde deel tot de conclusie dat Gilligan gelijk heeft met haar bewering dat de perspectieven van zorg en van rechtvaardigheid verschillende en niet tot elkaar te herleiden vormen van moreel denken impliceren. Identieke morele problemen krijgen een verschillende interpretatie, en ook de relevant geachte contextuele factoren zijn niet gelijk. Maar Gilligans pretentie dat beide benaderingen even adequaat zijn, kan niet staande gehouden worden. Ten eerste is de zorgbenadering niet goed toepasbaar op de (vele) ethische problemen die de sfeer van de persoonlijke verbondenheid van mensen overstijgen. Ten tweede kan, hoewel objectiviteit in de ethiek een altijd onbereikbaar ideaal zal blijven, het streven naar onpartijdigheid en algemeengeldigheid niet volledig worden opgegeven. In het licht van dat criterium is een zelfstandige zorg-ethiek niet overeind te houden - tenzij door haar zodanig bij te stellen dat ze in feite op die manier weer tot een variant van de theorie van Kohlberg is teruggebracht.

Het boek van Vreeke kan in verschillende opzichten een voorbeeldige theoretisch-pedagogische dissertatie worden genoemd. Niet in de zin van een briljant werk, maar, wat veel belangrijker is, omdat het de beheersing toont van het métier. Dat is een compliment zowel aan de auteur als ook aan zijn begeleiders. Als goede onderzoeksgroepen aan Nederlandse universiteiten vergeleken mogen worden met de plaatselijke gilden van weleer, waarin het handwerk op meestersniveau wordt beoefend en waarin daaromheen een aantal leerlingen (doctoraal studen-

ten) en gezellen (promovendi) het vak al doende leren, dan mag dit een proef-schrift heten in de beste zin van het woord. Dat begint bij de keuze en de afbakening van het onderwerp. Het onderzoeksproject betreft een interessante en belangrijke problematiek, die desalniettemin zonder verlies aan relevantie beperkt gehouden kon worden door de concentratie op een discussie tussen twee centrale auteurs. Het betoog is systematisch van opzet en duidelijk ingedeeld. Het bevat een zelfstandige en goed volgehouden probleemanalyse. De aangename dichtheid van argumentatie staat de helderheid niet in de weg. Tenslotte wordt het geheel gevoed door een goede keuze van literatuur.

Vreeke laat alle stappen zien van het theoretisch pedagogisch grondslagen-onderzoek. Na een korte plaatsbepaling van de onderhavige problematiek worden eerst de te behandelen standpunten zorgvuldig en gedetailleerd weergegeven, inclusief theoretische ontwikkelingen en wederzijdse punten van kritiek. Bijzondere aandacht wordt daarbij gegeven aan de impliciete en expliciete redeneringen die de beide visies schragen. Zo komen ook breuklijnen en theoretische contaminaties aan het licht. Van daaruit worden de grondslagen van het debat gereconstrueerd, waarbij tevens duidelijk wordt waarover het debat niet of slechts schijnbaar gaat. Op grond van een weloverwogen keuze en explicitering van criteria worden de verschillende pretenties tenslotte aan een beoordeling op hun relatieve houdbaarheid onderworpen. In een slotbeschouwing wordt, zij het nogal summier, de geboden analyse in haar bredere context teruggeplaatst en worden enige implicaties van het betoog aangegeven.

Dit alles in het bestek van minder dan 200 pagina's. Zelfs wie niet in het onderwerp geïnteresseerd is kan van dit boek veel leren.

A. W. van Haaften

H. Freudenthal

Revisiting mathematics education. Chinese lectures

Kluwer Academic Publishers, Dordrecht 1991 220 pagina's *f* 120,– ISBN 0 7923 1299 6

Freudenthal heeft voor zijn allerlaatste boek een opmerkelijke titel gekozen. Bij "Revisiting" denk je aan herziening en met "Revisiting Mathematics Education" kan dan bedoeld zijn dat de auteur het opportuun acht zijn ideeën over wiskundeonderwijs te herzien. Freudenthal heeft met deze titel echter iets anders op het oog namelijk een kritische terugblik op zijn werk, zijn ideeën en de wijze waarop die tot stand kwamen. Die ideeën worden niet zozeer herzien – al gebeurt dat ook wel – maar vooral bézien. Freudenthal onderwerpt fundamentele vraagstukken aan een grondige analyse en reflecteert daarbij op zijn eigen wetenschappelijk denken.

De ondertitel "China Lectures" is waarschijnlijk bedacht naar aanleiding van een studiereis die de auteur enkele jaren geleden naar China maakte.

Wie met het werk van Freudenthal al vertrouwd is, zal sommige hoofdstukken uit deze veelzijdige studie wat vlotter doorlezen, en lezers die voor het eerst een boek van Freudenthal ter hand nemen, ontmoeten een auteur met een eigen kijk op wiskunde, op onderwijs, onderwijsleerprocessen en onderwijsonderzoek.

In het leven van een mens begint wiskunde altijd als 'common sense', zo luidt een eerste grondgedachte, en het wiskundeonderwijs moet bij die 'common sense' aansluiten. "In general I believe that in instruction it would be more recommendable to start with common sense ideas rather than to reject them as outdated and better being suppressed" (p. 6). Het kunnen opzeggen van de getallenrij – het allereerste algoritme met een wiskundig karakter dat een kind verwerft – is een 'common sense' algoritme dat voortkomt uit 'means of common language'. Dat 'common sense' idee moet echter niet worden misverstaan. Er is niet mee bedoeld dat de kennisverwerving slechts op de empirie (zeg maar de dagelijkse ervaring) zou berusten. In de natuurkunde ontstaan concepties vaak uit ervaringen die kinderen en ook volwassenen spontaan beleven ('ijzer is kouder dan hout', 'verdampen dat is verdwijnen').

Het gaat om meer. Het wiskundeonderwijs moet weliswaar starten bij de 'common sense' ideeën, maar die ideeën moeten door bewustworden en reflectie op hoger niveau worden gebracht. Dat gebeurt via transformatie en formalisering. Om op hoger niveau weer 'common sense' te worden. Deze hogere niveaus worden niet bereikt als regels van bovenaf worden opgelegd: "But having be imposed they never had a real chance to develop into common

sense of a higher order" (p. 8). 'Common sense' treffen we dus aan, in elk geval in de wiskunde, op verschillende niveaus.

Het is dan ook niet toevallig dat de ontwikkeling van de wiskunde zich onderscheidt van bijvoorbeeld de ontwikkeling van de natuurwetenschap. De natuurkunde ontwikkelde zich sprongsgewijs en werd gekenmerkt door revoluties. De wiskunde evolueerde *tegelijkertijd* op verschillende plaatsen in verschillende culturen. "Common sense in order to become genuine mathematics and in order to progress, had to be systematised and organised" (p. 9). Met andere woorden de 'common sense' transformeerde in regels en die regels werden vervolgens weer 'common sense', maar nu van hogere orde. Modellen functioneren in dit proces als intermediair, ze ordenen de werkelijkheid en maken formalisering mogelijk.

'Guided reinvention' is een ander sleutelbegrip in de theorie van Freudenthal. Hij zet dit begrip graag af tegen - door hem als modieus ervaren - begrippen als 'discovery learning': 'uncovering what was covered up by somebody else - hidden eastern eggs'. Het gaat, met andere woorden meer om problemen die de onderzoeker in zijn hoofd had dan om de zoekprocessen van de lerende zelf. 'Reinvention' dus, maar wát moet dan uitgevonden worden, de wiskunde en de denkprocessen van onze voorouders? Nee, zegt Freudenthal, want het gaat om geleid *her*uitvinden. Waarheen moeten de kinderen geleid worden? Niet zozeer naar de wiskunde maar eerder naar het mathematiseren, abstraheren eerder dan abstracties, schematiseren eerder dan schema's, formaliseren eerder dan formules. Kortom, het gaat dus vooral om het (her)uitvinden van wiskundige activiteiten en processen. Hoe verloopt dat proces? Heruitvinden vraagt om interactieve instructie.

Freudenthal pleit krachtig voor nauwe banden met de realiteit. Wanneer de realiteit ontbreekt, raakt de zin van het mathematiseren volkomen zoek. De kinderen reproduceren dan weliswaar de juiste formules, maar kunnen die niet toepassen.

Uitvoerig gaat de auteur in op de vraag welke leerprocessen voor het wiskundeonderwijs van belang zijn. Hij bespreekt in dit verband onder meer de betekenis van het observeren van het handelen van anderen, zelfobserbatie, observatie van het leren in groepen, blikwisseling ('change of perspective') en het gedachtenexperiment. Met Van Hiele, naar wie hij ook verwijst, onderscheidt hij niveaus in het leer- en denkproces. Dit niveau-idee correspondeert met de opvatting van Freudenthal dat leerprocessen *discontinu* verlopen. Noodzakelijk voor niveauverhoging is volgens de auteur (die opvatting vinden we trouwens ook bij Piaget): "*Being made conscious and subjected to reflection*" (curs. van H.F). Reflectie stimuleert dus niveauverhoging: "...reflection on the higher level on one's activities of the lower one" (p. 101). Onder reflectie wordt in dit boek verstaan "...mirroring oneself in someone else in order to look through his skin, to explore him, to take him in" (p. 104). Reflectie is dus van oorsprong sociaal, zoals ook Vygotskij al zei. Het verplaatsen in een standpunt van een ander als zodánig rekent Freudenthal al tot reflectie. Men kan echter ook verdedigen dat er pas van reflectie sprake is als een persoon zijn handelen analyseert, als het ware vanuit het standpunt van de ander. Volgens deze opvatting bestaat reflectie uit de analyse van het *eigen* handelen.

Met algemene leertheorieën heeft de auteur niet veel op: "I strongly distrust general learning theories". Gal'perin is zo'n leerpsycholoog met wie hij graag de degens kruist. Volgens Gal'perin kan vrijwel alles gematerialiseerd worden en ook tekens zouden tot de materiële wereld behoren, zegt Freudenthal. Wellicht, maar Gal'perin zelf heeft de betekenis van tekens (en symboolgebruik) altijd erg betwijfeld en op dit punt zelfs Vygotskij bestreden. Bovendien is 'gematerialiseerd' en 'materieel' niet van dezelfde 'learning level' zoals Freudenthal zegt. Het gematerialiseerde bevat namelijk al niet-materiële, mentale elementen. Het probleem in de theorie van Gal'perin is veeleer dat op de instructie zoveel nadruk wordt gelegd dat de constructie door kinderen (zeg maar de 'common sense') te weinig ruimte krijgt.

Freudenthal is echter geen constructivist. Hij twijfelt er ernstig aan of de discussie, onder filosofen, over de vraag of en hoeverre een individu zijn eigen wereldbeeld construeert wel enige zin heeft. Hij geeft de voorkeur aan het begrip *re*constructie. Ik vraag me af of het allemaal zoveel uitmaakt. In wezen gaat het er toch om dat een kind ruimte wordt geboden zijn eigen ideeën (constructies) te vormen. Door interactie worden die uitgelokt en door reflectie op hoger niveau gebracht.

Freudenthal heeft een uitgesproken visie op onderwijsonderzoek en in die visie staat het begrip 'validiteit' centraal. Het streven via steeds geavanceerder wiskundig-statistische technieken de betrouwbaarheid te verhogen, gaat vaak ten koste van de validiteit van onderzoek. De meetprocedures zijn weliswaar uiterst verfijnd, maar de onderzoekers weten soms nauwelijks wat ze nou eigenlijk onderzoeken, omdat ze zich onvoldoende voorbereid wagen op het gebied van het wiskundeonderwijs. Het misbruik van wiskundige procedures is de auteur een doorn in het oog evenals misinterpretaties van onderzoeksgegevens. Zo zou uit een 'otherwise excellent national' onderzoek (bedoeld is hier het P.P.O.N. onderzoek van het C.I.T.O.) zijn gebleken dat meisjes op reken/wiskundegebied zwakker presteren dan jongens. Deze interpretatie van de onderzoeksgegevens kan niet juist zijn, aldus Freudenthal, omdat zoals bekend de jongens 'the overwhelming majority' vormen van de kinderen die in Nederland in het speciaal onderwijs belanden, zodat de besten dus overblijven in het basisonderwijs. Onlangs echter, in een zogenoemde mediopeiling van het C.I.T.O., is gebleken dat de meisjes beter dan jongens presteren bij het 'begrijpend lezen' van zakelijke teksten. Worden jongens dan alleen op grond van hun tegenvallende rekenprestaties naar het speciaal onderwijs verwezen? Deze bevinding lijkt op gespannen voet te staan met de verklaring van Freudenthal.

"Revisiting Mathematics Education" is een rijk boek waarin maar weinig thema's die het wiskundeonderwijs raken, onbesproken blijven. De auteur zal - mede door zijn vrijmoedige en directe betoogtrant - bij sommige lezers nog wel eens tegenspraak en discussie oproepen. Dit zal de fundamentele koerswijziging in het denken over wiskundeonderwijs, waarvoor Freudenthal in zijn werk steeds heeft gepleit, in elk geval niet meer kunnen verhinderen.

J. M. C. Nelissen

T. T. D. Peetsma

Toekomst als motor? Toekomstperspectieven van leerlingen in het voortgezet onderwijs en hun inzet voor school

Academisch proefschrift Universiteit van Amsterdam.
SCO, Amsterdam 1992, 181 pagina's
ISBN 90 681 3317 9

Peetsma heeft een onderzoek verricht naar de rol die het toekomstperspectief als motivationele variabele speelt tijdens de schoolloopbaan. Zij kiest als theoretische basis niet voor de sociale-leertheorieën, waarin de neiging om naar toekomstige doelen te streven afhankelijk wordt gesteld van de verwachting dat daarmee een beloning wordt verkregen én van de verwachte waarde van die beloning. Zij kiest voor de interactionistische benadering, waarin de verklaring van gedrag vooral wordt gezocht in een wisselwerking tussen persoonlijkheidstrekken en situatiekenmerken. Kenmerkend voor deze benadering is volgens Peetsma de nadruk die wordt gelegd op afwegingsprocessen die uitmonden in intentioneel gedrag. Echter, ook bij sociale-leertheorie hangt de handeling af van de afweging die iemand maakt m.b.t. de waarde die subjectief wordt gehecht aan het doel en de even subjectieve verwachting dat doel te kunnen bereiken met behulp van bepaalde gedragskeuzen. In het theoretische hoofdstuk wordt nauwelijks duidelijk gemaakt waarom bepaalde benaderingen worden afgewezen en een andere wordt gekozen. Bij de onderzoeksopzet vinden we er dan ook weinig van terug. Zeker niet de voor de interactionistische benadering zo kenmerkend geachte afwegingsprocessen.

De korte theoretische uiteenzetting wordt gevolgd door een beknopt overzicht over de uitkomsten van allerlei onderzoekingen op het terrein van haar probleemstelling. Ondanks de beknoptheid worden er ook uitkomsten vermeld van voor deze studie in het geheel niet relevante onderzoekingen op het gebied van de psychopathologie. Peetsma stelt overigens dat het vele onderzoek weinig overeenstemmende resultaten heeft opgeleverd. De lezer is dan geneigd te verzuchten: "dat heb ik vaker gelezen".

Toch vallen er wel enkele patronen te onderkennen. Ik zal er een paar noemen die in het kader van deze studie wel relevant lijken, en ook weer door deze studie worden bevestigd. Jongens zijn meer gericht op een beroepsloopbaan

en meisjes meer op sociale relaties. Recenter onderzoek toont aan dat er verschuivingen optreden. Meisjes worden moderner. Zij richten zich nu op een 'dubbelleven' waarin sociale relaties, of gezinsvorming, wordt gecombineerd met beroepsuitoefening. Ook het perspectief van jongens wordt breder; ze richten zich voor later niet meer alleen op een beroep maar ook op hun vrije-tijdsbesteding. Helaas zal die ontwikkeling het 'dubbelleven' van vrouwen niet verlichten.

In dit onderzoek wordt toekomst*perspectief* omschreven als een attitude van iemand ten aanzien van een doelobject op een zekere termijn in de toekomst. Op pagina 107 schrijft zij zelfs: "Toekomstperspectief is opgevat als een motivationele variabele". Als dit begrip dus wordt opgevat als iets dat voor een *houding* of *motivatie* staat, dan is de gekozen term niet erg gelukkig. Toekomst*gerichtheid* ware dan te verkiezen. Zoals Peetsma ook zelf aangeeft (p. 30) kan men bij attituden een cognitief, een affectief en een intentioneel element onderscheiden. Naar mijn smaak zou men het cognitieve element heel adequaat als het toekomstperspectief kunnen omschrijven. Hoe ziet men de toekomst in het algemeen en zijn/haar toekomstkansen in het bijzonder? Denkt men inderdaad ver vooruit of niet? Bij het gevoelsmatige aspect kan men nagaan of iemand beducht is voor de toekomst of die juist optimistisch en vol verwachting tegemoet treedt? Ten slotte het gedragsintentioneel aspect. Naar dat laatste is Peetsma op zoek. Hebben leerlingen met een bepaalde toekomstgerichtheid ook een bepaalde gedragsintentie, en zo ja, komt die ook tot uiting in hun inzet voor school?

Op grond van haar eerdere onderzoek komt Peetsma tot de keuze van vragenlijstonderzoek met behulp van schalen. Haar argumenten voor deze keuze zijn overtuigend. De betrouwbaarheid en de geldigheid van de tot nu toe gebruikte en door Peetsma besproken alternatieve (indirecte) methoden maken absoluut geen toereikende indruk. De door haarzelf ontwikkelde schalen blijken bijna allemaal zeer betrouwbaar. Een puntje van kritiek kan worden geplaatst bij de gebruikte antwoordcategorieën. In plaats van het gebruikelijke neutrale antwoord is gekozen voor "past een beetje bij mij". Dit heeft nog altijd een positieve lading. Ook de beide naastliggende antwoordcategorieën zijn ongebruikelijk verwoord: "past vrij goed bij mij" en "past nauwelijks bij mij". Waarom niet gewoon: heel goed, goed, niet goed/niet slecht, slecht en heel slecht, ondersteund door symbolen die middelbare scholieren heel goed begrijpen zoals ++, +, +/-, -, --?

De uitkomsten geven het volgende beeld. Het perspectief op school- en beroepsloopbaan op lange termijn wordt positiever naarmate de leerlingen een hoger schooltype bezoeken. Het korte termijn perspectief wordt minder positief naarmate het leerjaar stijgt. De examens werpen kennelijk steeds duidelijker hun dreigende schaduwen vooruit. De inzet voor school neemt af naarmate het leerjaar stijgt. Dit lijkt mij een tamelijk verrassende uitkomst. Algemeen wordt aangenomen dat de leerlingen in de brugklas heel positief aan hun loopbaan in het voortgezet onderwijs beginnen. Daarna neemt de inzet af. Het nieuwtje is er af, andere interesses eisen de aandacht op. Pas tegen de tijd dat de examens in zicht komen, met name in de examenklas zelf, zal de inzet weer sterk gaan toenemen. Kortom, een kromlijnig verband ligt voor de hand.

Het onderzoek toont een verband aan tussen gemotiveerdheid en de toekomstperspectieven op school- en beroepsloopbaan, en tussen gemotiveerdheid en twee andere toekomstperspectieven die werden onderscheiden: sociale relaties en eigen ontwikkeling. Het perspectief op vrije tijd lijkt echter averechts te werken voor de inzet op school.

Het onderzoeksdeel van de studie toont grote degelijkheid, al zien de in de bijlage afgedrukte vragenlijsten er als drukwerk weinig professioneel uit. De degelijkheid schuilt voor een groot deel in het feit dat Peetsma kan voortbouwen op eerder door haar uitgevoerde studies op dit terrein. Van de statistische analyses in vele soorten is veel werk gemaakt. De verslaggeving daarvan is echter gortdroog: voer voor ingewijden. Ook in het slothoofdstuk over de uitkomsten blijft de discussie zeer dicht in de buurt van de verkregen uitkomsten. Daardoor is deze helder geschreven dissertatie helaas weinig inspirerend, maar dat laatste vormt geen relevant criterium voor een promotie.

M. J. de Jong

Mededelingen

Symposium Persoonlijkheid: ontwikkelingspsychologie en ontwikkelingspsychopathologie

Het thema van dit symposium, dat op 10 en 11 juni 1993 te Leiden wordt gehouden, is de relatie tussen persoonlijkheidskenmerken en psychopathologie bij kinderen en adolescenten. De organisatie berust bij de PAO, studierichting Gedragswetenschappen, het Academisch Centrum Kinder- en Jeugdpsychiatrie Curium, en de Secties Kinder- en Jeugdpsychiatrie van de Rijksuniversiteit Leiden en de Nederlandse Vereniging voor Psychiatrie.

Inschrijvingen en inlichtingen bij: PAO, Postbus 325, 2300 AH Leiden, tel.: 071-143142, fax: 071-140145.

Congres: Onderwijs en Stress

Op 7 oktober 1993 organiseert het Stress en Traumafonds een ééndaags congres over de problemen van stress in het onderwijs, zowel bij leraren als bij leerlingen.

Nadere informatie is verkrijgbaar bij het Congresbureau Van Namen & Westerlaken, Postbus 1558, 6501 BN Nijmegen, tel.: 080-234471, fax: 080-601159.

Inhoud andere tijdschriften

Comenius
12e jaargang, nr. 4, 1992

Thema: Vygotksky's legacy
Vygotksky's Legacy: Understanding and Beyond. An introduction, door S. Miedema, M.H. van IJzendoorn en R. van der Veer
Children's contribution to their development as a theme in Vygotsky's work, door E. Elbers
Vygotsky's dynamic systems, door P. van Geert
Vygotsky's view on schooling. Reflections on R. van der Veer and J. Valsiner, Understanding Vygotsky. A quest for synthesis, door W. Wardekker
Vygotsky's modern appeal, door S. Bem
Voices at play. Understanding Van der Veer and Valsinger, door R. van der Veer en J. Valsiner
Jongenswereld - meisjeswereld? Jongeren over de relatie met hun ouders en leeftijdgenoten, door J. de Hart

Pedagogisch Tijdschrift
17e jaargang, nr. 5/6, 1992

Thema: Wetenschapsgeschiedenis van de pedagogiek
Historische processen in de ontwikkeling van het experimenteel onderzoek in de Pedagogische Wetenschappen, door M. Depaepe
Van natuurkundige naar biologische pedagogiek. De paradigmatische omslag tussen circa 1890 en 1923, door P. Halens
Eugenetisch pessimisme en pedagogisch optimisme bij J. Demoor.
Case-study naar het geneeskundig denken over abnormaliteit eind negentiende - begin twintigste eeuw, door Ph. de Wilde
Voorgeschiedenis van de Nederlandse orthopedagogiek in verband met de ontplooiing van het speciaal onderwijs, gedurende de eerste helft van de twintigste eeuw, door J.C. Sturm
Kohnstamm en het idee van kinderlijke ontwikkeling. Pedagogiek en psychologie, door E. Mulder
Feiten en waarden in de ontwikkeling van de Nederlandse academische pedagogiek. Een opmaat voor pedagogisch wetenschapsonderzoek, door G. Biesta en S. Miedema
'De Zijlijn'. Vrouwen en de pedagogische wetenschap in Nederland tussen 1779 en 1946, door M. van Essen en M. Lunenberg

Tijdschrift voor Onderwijsresearch
17e jaargang, nr. 4, 1992

Stand van zaken ontwikkelingsproces (voortgezet) speciaal onderwijs: constructie van een instrument, door E. Kool, D. Rooth, C. Bruinsma en A. van der Leij
De GIVON, de getrapte indeling naar voltooid onderwijsniveau, door W. Jansen Heijtmajer
Effect van tekststructuur op de reproduktie van informatieve teksten bij brugklasleerlingen, door M.J.M. Voeten, A. Bergmans, C.A.J. Aarnoutse en I. de Groot
Notities en Commentaren
Het beoordelen van de SVO-beoordeling, door P. Vedder

Tijdschrift voor Onderwijsresearch
17e jaargang, nr. 5, 1992

Thema: De relatie tussen cognitieve en metacognitieve vaardigheden
De relatie tussen cognitieve en metacognitieve vaardigheden: een introductie, door M. V. J. Veenman, A. Samarapungavan, B. H. A. M. van Hout-Wolters en J. J. Beishuizen
Metacognitie van lerenden in onderwijsleerprocessen, door M. Elshout-Mohr
Intelligentie en metacognitieve vaardigheden, door M. V. J. Veenman en J. J. Elshout
The relation between metacognitive and domain specific knowledge, door A. Samarapungavan en M. Milikowski
Cognitie en metacognitie bij het bestuderen van informatieve tekst, door E. Stoutjesdijk en J. J. Beishuizen

Tijdschrift voor Orthopedagogiek
31e jaargang, nr. 12, 1992

'De wolken waarom moeder schreide'. Oratie van D. J. de Levita
Wachter wat is er van de nacht? Kanttekeningen bij de inaugurele rede van prof.dr. D. J. de Levita, door J. van Weelden
Gezinsproblemen in de welwaarts- en verzorgingsstaat, door J. van Acker
Stress bij ouders van een verstandelijk gehandicapt kind. Evaluatie van thuiszorg, door H. W. van Berkum

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 1, 1993

Primaire preventie van kindermishandeling, door C. Hoefnagels
Oordeelsvorming en remediëringskansen in het speciaal onderwijs, door J. Th. A. Bakker en C. B. L. M. Ubachs
Het Leerlingvolgsysteem van het Cito; inhoud en werkwijze, door P. Gillijns
Reactie. Vroegbehandeling en het jonge lichamelijk gehandicapte kind: een reactie uit de praktijk, door J. Hendriksen en A. Reinders

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 2, 1993

Thema: Zorgen voor de leerling
Zorg om de verzorging door R. de Groot
Onderwijsverzorging: gevierendeeld tussen overheid, scholen, markt en vooral zichzelf, door K. B. Koster
Leerlingbegeleiding na 1995. De WRR ziet vakinhoudelijke ontwikkeling over het hoofd, door W. Meijer
De toekomst van de onderwijsbegeleiding en WSNS, door R. J. Gorter
De structuur en organisatie van de begeleiding van de sociaal-emotionele ontwikkeling van leerlingen, door C. F. M. van Lieshout en G. J. T. Haselager

Ontvangen boeken

Aarssen, J., Ruiter, J. J. de, & Verhoeven, L, *Toetsing Turks en Arabisch aan het einde van het basisonderwijs*. Tilburg University Press, Tilburg, 1992, *f* 29,90.

Daal, V. H. P. van, *Computer-based reading and spelling practice for young dyslexics* (dissertatie). Vrije Universiteit, Amsterdam, 1993.

Eeden, P. van den, Jong, U. de, Koopman, P., & Roeleveld, J., *Schoolloopbanen in Amsterdam. Leerlingen op Amsterdamsche scholen voor algemeen voortgezet onderwijs in de jaren zeventig* (Forum-reeks 15). Swets & Zeitlinger, Lisse, 1993, *f* 42,–.

Hamers, J. H. M., Sijtsma, K., & Ruijssenaars, A. J. J. M. (Eds.), *Learning potential assessment. Theoretical, methodological and practical issues*. Swets & Zeitlinger, Lisse, 1993, *f* 98,–.

Hofman, R. H., *Effectief schoolbestuur. Een studie naar de bijdrage van schoolbesturen aan de effectiviteit van basisscholen* (dissertatie). RION, Groningen, 1993.

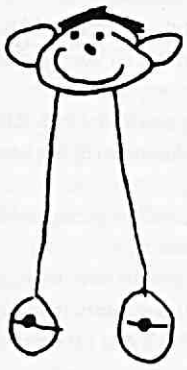
Linden, W. J. van der, Mommers, M. J. C., Peters, J. J., Tomic, W. (red.), *Onderwijskunde in de klas*. Swets & Zeitlinger, Lisse, 1993, *f* 38,–.

Nijhof, W. J., Franssen, H. A. M., Hoeben, W. Th. J. G., Wolbert, R. G. M. (red.), *Handboek curriculum. Modellen, theorieën, technologieën*. Swets & Zeitlinger, Lisse, 1993, *f* 79,50.

Oers, B. van, & Janssen-Vos, F. (red.), *Visies op onderwijs aan jonge kinderen*. Van Gorcum, Assen, 1992, *f* 27,50.

Swets & Zeitlinger Publishers

Heereweg 347, 2161 CA Lisse



Tekenen – een natuurlijke bezigheid

De psychologie van kindertekeningen

Glyn V. Thomas en Angèle M.J. Silk

ISBN 90 265 1318 6 180 PAG. *f* 29,50

Het meest recente en uitgebreide overzicht van psychologische theorieën over kindertekeningen. Waarom tekent een kind? Hoe tekent een kind? Wat zegt een tekening over ontwikkeling, niveau, begaafdheid, originaliteit, emoties, expressiviteit? Vanuit de ontwikkelingspsychologie, de klinische psychologie, de artistieke benadering en de cognitieve psychologie kijkt men naar verschillende aspecten. Recent onderzoek is vooral gericht op de analyse van het tekenproces en op het vaststellen van de informatie die kinderen in hun tekeningen verwerken.
Vertaling van *An Introduction to the Psychology of Children's Drawings*.
De auteurs zijn als psycholoog verbonden aan de Universiteit van Birmingham (Engeland).

Opvoeding en opvoedingsproblemen in het gezinsleven

P.M. Schoorl

ISBN 90 265 1265 1 184 PAG. *f* 45,—

In elk gezin ontstaan vanzelfsprekende handelingspatronen, waarvan de gezinsleden zich niet altijd bewust zijn. Dit boek geeft een handleiding voor het blootleggen van deze impliciete gezinsregels in probleemsituaties, als noodzakelijke voorbereiding op de hulpverlening. De basis is een theoretische analyse van de regels in het gezinsleven en hun invloed op het functioneren van het gezin. Het resultaat is een praktisch en overzichtelijk boek over gezinsopvoeding, dat door de heldere en systematische opbouw ook zeer geschikt is voor opleidingen.

Verkrijgbaar in de boekhandel of bel: 02521 - 35111

15 Oct 6441

Inhoudsopgave

Pedagogische Studiën Jaargang 70, nummer 2/1993

G.W. Meijnen en J. Kloprogge
De evaluatie van het Onderwijsvoorrangsbeleid
Een introductie op de artikelen 82

G.W. Meijnen en J. Kloprogge
De evaluatie van het Onderwijsvoorrangsbeleid geëvalueerd 84

M.P.C. van der Werf, P. Tesser en L. Mulder
De prestaties van de OVB-doelgroepen in 1988: stand van zaken na de eerste meting 97

M.P.C. van der Werf en M.G. Weide
Effectieve voorrangskenmerken in de school en in de klas 108

P. Tesser en H. Vierke
De prestaties van de OVB-doelgroepen in 1990: stand van zaken na de tweede meting 122

K.M. Stokking
Differentiatie en zorgverbreding in het basisonderwijs: de behandeling van een vernieuwingsonderwerp in onderzoek en beleid, 1984-1990 135

Boekbesprekingen 151

Mededelingen 159

De evaluatie van het Onderwijsvoorrangsbeleid

Een introductie op de artikelen

G. W. Meijnen en J. Kloprogge

Het Onderwijsvoorrangsbeleid (OVB) is in 1983 gestart als vervolg op het Stimuleringsbeleid. Jaarlijks is met dit beleid meer dan 400 miljoen gulden gemoeid. Met de start ervan is tevens door de overheid het voornemen uitgesproken dit langdurige beleidsprogramma te doen evalueren. Dit voornemen is in daden omgezet en er loopt sedert 1988 een grootschalig evaluatieprogramma. De start van het eigenlijke evaluatie-onderzoek is vooraf gegaan door een voorbereidingsperiode die enkele jaren in beslag heeft genomen. Inmiddels zijn er tal van onderzoeksverslagen verschenen, voornamelijk in de vorm van rapporten en terugrapportages aan betrokkenen. De tijd is echter nu ook rijp voor een presentatie van resultaten die een beeld schetsen van de kern van wat in de afgelopen jaren is onderzocht.

Meijnen en Kloprogge beschrijven in hun openingsartikel: *De evaluatie van het onderwijsvoorrangsbeleid geëvalueerd*, hoe het evaluatieprogramma tot stand is gekomen en hoe het tot nu toe verloopt. In het bijzonder wordt daarbij gerefereerd aan die onderwerpen die in de internationale literatuur over programmaevaluatie als cruciaal worden beschouwd. Een evaluatie van het evaluatie-onderzoek derhalve.

Van der Werf, Mulder en Tesser presenteren vervolgens in hun artikel: *De prestaties van de OVB-doelgroepen in 1988: stand van zaken na de eerste meting*, gegevens over de taal- en rekenprestaties van diverse groepen leerlingen in groep 4, 6 en 8. Voorts rapporteren zij over de prestaties gescoord op een intelligentietest en over de relaties tussen enkele gezinskenmerken en de prestaties van de leerlingen. In de presentatie maken zij telkenmale onderscheid tussen 1.90 leerlingen (allochtone leerlingen uit achterstandssituaties), 1.25 leerlingen (autochtone leerlingen uit achterstandssituaties) en 1.00 leerlingen (geen achterstandssituatie).

In het derde artikel, van de hand van Van der Werf en Weide, getiteld: *Effectieve voorrangskenmerken in de school en in de klas*, wordt nagegaan welke school- en klaskenmerken, differentieel van invloed zijn op de leerprestaties van de leerlingen. In het bijzonder wordt daarbij een onderscheid aangebracht tussen scholen die helemaal geen OVB-faciliteiten genieten en scholen die zowel formatie- als gebiedsfaciliteiten dan wel alleen formatiefaciliteiten ontvangen.

Het vierde artikel, geschreven door Tesser en Vierke, getiteld: *De prestaties van de OVB-doelgroepleerlingen in 1990: stand van zaken na de tweede meting*, behelst de resultaten van de eerste herhaling van de cohortmetingen. De leerlingen uit groep 4 in 1988 zijn nu, in het merendeel van de gevallen, doorgestroomd naar groep 6 en die van groep 6 naar groep 8. Groep 4 in 1990 is nieuw. Deze gegevens openen de mogelijkheden voor het analyseren van ontwikkelingen bij leerlingen in de tijd. De auteurs rapporteren daarover ook enige gegevens. Daarbij wordt onderscheid gemaakt naar leerlingen met uiteenlopende gewichten en naar scholen die geen dan wel een of meerdere typen faciliteiten ontvangen. Ten slotte berekenen de auteurs of een groot percentage leerlingen uit achterstandssituaties op school op zich een ongunstige factor vormt voor de prestatie-ontwikkeling van de leerlingen op die school.

Het vijfde artikel is geschreven door Mulder en draagt als titel: *De invloed van het advies op de schoolloopbaan van de OVB-doelgroepleerlingen in het voortgezet onderwijs*. De afgelopen jaren hebben diverse onderzoekers melding gemaakt van hun bevinding dat er bij adviezen aan allochtone leerlingen bij de overgang van basis- naar voortgezet onderwijs mogelijk lagere eisen worden gesteld dan aan die bij autochtone leerlingen. Mulder toetst deze veronderstelling en gaat vervolgens na wat de relatie is tussen de gegeven adviezen en de onderwijspositie van leerlingen na twee jaar voortgezet onderwijs. Uiteraard maakt zij bij

deze analyse onderscheid tussen leerlingen met uiteenlopende leerlinggewichten.

Het zesde artikel, waarvan Driessen de auteur is, is niet gebaseerd op gegevens die verzameld zijn in het kader van het cohortonderzoek dat zijn wortels heeft in het basisonderwijs, maar op gegevens verkregen door een cohortonderzoek van het CBS in het eerste leerjaar van het voortgezet onderwijs in 1989 (VOCL). Dit cohort moet zowel gegevens opleveren voor de evaluatie van de Basisvorming als van het Onderwijsvoorrangsbeleid. Hoewel in Onderwijsvoorrangsgebieden ook scholen voor voortgezet onderwijs in het samenwerkingsverband zijn opgenomen, is over hun leerlingenpopulatie weinig bekend. Alle aandacht concentreert zich op de basisscholen binnen de gebieden. Driessen richt zich in zijn artikel dan ook op deze scholen. De titel van zijn bijdrage luidt dan ook: *Binnen en buiten onderwijsvoorrangsgebieden. Sociaal-etnische achtergronden en onderwijspositie van leerlingen in het voortgezet onderwijs.*

Het zevende en laatste artikel: *Samenwerking in onderwijsvoorrangsgebieden*, is van de hand van Verhaak en behandelt de problematiek van de samenwerking tussen diverse instellingen binnen onderwijsvoorrangsgebieden. Deze gebieden worden door de overheid gefaciliteerd met de bedoeling in onderlinge samenwerking de problematiek van de onderwijsachterstanden te bestrijden. Basisscholen kunnen samenwerken met bibliotheken of scholen voor speciaal onderwijs dan wel met scholen voor voortgezet onderwijs of andere instellingen in de wijk. Verhaak beschrijft of er wordt samengewerkt, op welke onderdelen en wat de belemmeringen zijn bij pogingen tot samenwerking.

Het vijfde tot en met het zevende artikel zullen niet in dit nummer (2) maar in nummer 4 van deze jaargang verschijnen.

De evaluatie van het Onderwijsvoorrangsbeleid geëvalueerd

G. W. Meijnen en J. Kloprogge

Samenvatting

In dit artikel wordt het Landelijk Evaluatie-Onderzoek van het Onderwijsvoorrangsbeleid (LEO-OVB) geanalyseerd (evaluatie van een evaluatie). Zowel de voorbereidingsfase als de uitvoering van het onderzoek worden beschreven. In het artikel wordt bovendien ingegaan op die onderwerpen die in de internationale literatuur als de meest relevante worden beschouwd in het kader van programma-evaluatie. Tot die onderwerpen behoren: designkeuze, de verhouding tussen interne en externe validiteit en de relaties met de omgeving van het evaluatieprogramma. De conclusie luidt dat zowel het ontwerp als de verbindingen met de context nog steeds beantwoorden aan recente inzichten omtrent opzet en inrichting van een extern evaluatie-onderzoek. Tevens wordt opgemerkt dat dit onderzoeksprogramma niet het totale Onderwijsvoorrangsbeleid evalueert.

1 Inleiding

In 1991 verscheen het Negende Jaarboek van de National Society for the Study of Education (McLaughlin & Phillips (Eds.), 1991). Een aantal gerenommeerde, deels bejaarde, evaluatiedeskundigen schetst daarin de ontwikkelingen in hun eigen denken over evaluatie-onderzoek gedurende de laatste vijfentwintig jaar. De bijdragen concentreren zich in het bijzonder op de problematiek van programma-evaluatie. Opvallend is de convergentie in denken omtrent wat de belangrijkste onderwerpen zijn waar het programma-evaluatie betreft. Het betreft dan in het bijzonder (1) de notie dat programma-evaluatie wordt uitgevoerd binnen een maatschappelijk krachtenveld, (2) de spanning die bij de opzet van het programma optreedt in termen van interne en externe validiteit en, derhalve, (3) wat de meest geëigende designs zijn voor programma-evaluatie, (4) wat verwacht mag worden omtrent de impact van de verkregen resultaten op de beleidsvoering en (5) de wetenschapsinterne ontwikkelingen in het evaluatie-onderzoek. In dit artikel zal het in uitvoering zijnde Landelijk Evaluatie-Onderzoek Onderwijsvoorrangsbeleid (LEO-OVB) mede in het perspectief van deze onderwerpen worden geanalyseerd.

Nadat zowel een beschrijving van de voorbereiding van het programma is gegeven alsook een beschrijving van het functioneren van 1986 tot 1992, worden deze periodes geëvalueerd waarbij deze onderwerpen onder meer aan de orde zullen komen. Ook deze activiteiten kunnen als een poging tot evaluatie aangemerkt worden (evaluatie van de evaluatie). Een van de punten van consensus onder de deskundigen is dat een externe instantie, dan wel een niet gecommitteerde onderzoeker de eerst aangewezene is om deze taak op zich te nemen. Uit de rest van het betoog zal duidelijk worden dat de positie van de auteurs hier een hachelijke is, ze zijn beiden, hetzij als ontwerper, hetzij als programmacoördinator, betrokken bij het LEO-OVB.

2 De voorbereiding van het LEO-OVB

In 1983 verzoekt het Ministerie van Onderwijs en Wetenschappen een drietal onderzoekers van de Rijksuniversiteit Groningen (Hofstee, Koster, Meijnen) een plan te ontwerpen voor de evaluatie van het Onderwijsvoorrangsbeleid waarvan de contouren waren vastgelegd in een concept Onderwijsvoorrangsplan (definitieve versie: 1985) dat verscheen onder de verantwoordelijkheid van de toenmalige minister Van Kemenade en staatssecretaris Deetman. Een daarop gebaseerd wetsontwerp zou kort daarop bij de Tweede Kamer worden ingediend. In het toentertijd circulerende concept van de Onderwijsvoorrangswet (1984) wordt als centrale doelstelling van het beleid vermeld: "Het onderwijssvoorrangsbeleid is er op

gericht (...) *onderwijsachterstanden* van leerlingen ten gevolge van sociale, economische en culturele omstandigheden op te heffen of te verminderen". In de Memorie van Toelichting wordt onder 'onderwijsachterstand' verstaan; "(...) achterstand in ontwikkeling en in onderwijsleerprestaties in ruime zin".

De overheid tracht deze doelstelling te bereiken door (a) scholen met leerlingen uit achterstandssituaties extra te faciliteren (de zogenaamde gewichtenregeling) en door (b) faciliteiten ter beschikking te stellen aan *gebieden* met een concentratie van deze scholen.

Het verzoek aan de drie onderzoekers kan worden beschouwd als een hernieuwde poging van de overheid om een beleidsprogramma op dit terrein te evalueren. De eerste poging vond plaats in de periode 1977-1981. In dat laatste jaar verscheen namelijk het eindrapport van de evaluatie van het Stimuleringsbeleid, met min of meer soortgelijke doelstellingen, dat in 1974 was gestart (Mens & Van Calcar, 1981). Het overgrote deel van het verslag behelst beschrijvingen van activiteiten die in het veld plaatsvonden zowel op scholen als binnen betrokken welzijnsinstellingen, waarbij ook bestuurlijk-organisatorische aspecten niet uit het oog werden verloren. Van de oorspronkelijk geplande effectevaluatie werd zo goed als niets gerealiseerd. In evaluatiejargon: het onderzoeksverslag behelsde bijna uitsluitend procesevaluatie. In het Ten Geleide bij het verslag somt de Begeleidingscommissie een aantal problemen op die de uitvoering van de evaluatie hebben bemoeilijkt en voorts kan men zinsneden van het volgende type lezen: "Een gedetailleerde bespreking, laat staan een beoordeling ligt niet op de weg van de begeleidingscommissie" en: "Wordt nu de inspanning en andere investeringen van alle betrokkenen in dit project (het evaluatie-onderzoek) gerechtvaardigd door het resultaat?"

Uit het Ten Geleide en de overige verslagen kan men opmaken dat de meeste problemen zijn voortgekomen uit de geringe consensus omtrent de te evalueren doelstellingen, door de discussie omtrent de vraag of sommige doelstellingen wel gemeten zouden kunnen worden en uit de variatie in visies op de vraag wiens belangen de evaluatie zou moeten dienen (scholen, wetenschap, overheid). De onderzoekers die hierover zo ook hun eigen soms zeer geprononceerde, maar tevens uiteenlopende, meningen hadden, speelden zowel de rol van participant in de doelstellingendiscussie als uitvoerder van het evaluatie-onderzoek. Uit de verslagen wordt duidelijk dat de betrokken onderzoeksgroepen zowel van eigen programmadoelen als van eigen evaluatieopvattingen zijn uitgegaan. De Begeleidingscommissie is kennelijk niet in staat geweest er een geheel van te smeden. Dit verschijnsel had zich intussen ook bij de evaluatie van andere beleidsprogramma's voorgedaan (Scheerens, 1983). Het Ministerie van Onderwijs en Wetenschappen wilde dit probleem bij de evaluatie van het in gang te zetten Onderwijsvoorrangsbeleid voorkomen door een evaluatieplan te laten ontwerpen door deskundigen die niet op voorhand ook de uitvoering ter hand zouden moeten nemen.

Gelet op de ruim geformuleerde doelstellingen door het beleid hield de keuze van de aan te zoeken deskundigen ook een sturing in van de specificering van de doelstellingen. De jaren zeventig werden immers gekenmerkt door het debat tussen voorstanders van de zogenaamde 'brede doelstellingen' voor het Stimulerings/Onderwijsvoorrangsbeleid (cognitieve, sociaal-emotionele, expressieve) en de aanhangers van het beleid gericht op 'enge (cognitieve) doelstellingen' (schoolloopbanen, leerprestaties met betrekking tot de kernvakken). Dit debat was in 1983 nog niet geëindigd hoewel reeds een kentering merkbaar was ten faveure van de laatstgenoemde richting. De keuze van de deskundigen voor het ontwikkelen van het evaluatieplan was in dit kader niet oninteressant gegeven de beperkte middelen voor onderzoek. De te kiezen prioriteiten in het evaluatieonderzoek zouden in de richting van de ene of de andere doelstellingenfilosofie moeten uitvallen. Het verzoek aan de drie Groningse wetenschappers was in dit bestek een duidelijke: alle drie hadden zich voorstander betoond van de accentuering van de cognitieve doelstellingen in het onderwijs.

De onderzoekers hebben het onderzoeksprogramma vervolgens geschreven na consultatie van een groot aantal deskundigen en na overleg met een interdepartementale werkgroep van ambtenaren, belast met de voorbereiding van de beleidsevaluatie van het onderwijsvoorrangsplan.

Hofstee, Koster en Meijnen (1984) schrijven in een paragraaf voorafgaande aan de eigenlijke opzet van het onderzoek, dat bestudering van eerder verricht evaluatie-onderzoek naar landelijke innovatieprogramma's hen tot de volgende conclusies heeft gebracht (p. 9). "Ten eerste: het is noodzakelijk dat bij de start van een evaluatie-onderzoek de beleidsdoelstellingen van een vernieuwingsprogramma voldoende worden geëxpliciteerd. Ten tweede: na de vaststelling van een evaluatieplan in hoofdlijnen, dient de uitvoerende groep van onderzoekers daaraan relatief onafhankelijk uitvoering te kunnen geven en dient ook voor het overige voldaan te zijn aan minimale voorwaarden voor onafhankelijk onderzoek. Vierjaarlijkse bijstelling van het evaluatieplan lijkt een goede mogelijkheid om ook de voortgang van de evaluatie te beoordelen. Ten derde: de ervaringen met de evaluatie van grootscheepse onderwijsvernieuwingsprogramma's tonen aan dat variabelen op klasse- en leerlingniveau in het onderzoek worden verwaarloosd. Ten vierde: de uitkomsten van evaluatie-onderzoek dienen bij voorkeur een plaats te krijgen in de procedure voor de vaststelling van de onderwijsvoorrangsplannen".

Met betrekking tot de randvoorwaarden en uitvoering van het evaluatieonderzoek merken Hofstee e.a. ten slotte nog het volgende op. Zij wensen uitdrukkelijk onderscheid te maken tussen *beleidsevaluatie* en *extern evaluatie-onderzoek*. Beleidsevaluatie omvat meer dan effectonderzoek. Beleidsevaluatie dient bijvoorbeeld rekening te houden met gevestigde belangen, met maatschappelijke en politieke acceptatie van beleid, met werkgelegenheidsproblemen en dergelijke. Het voorgestelde onderzoek wordt door hen dan ook geen beleidsevaluatie genoemd maar een extern evaluatie-onderzoek. De beleidsevaluatie zou in hun ogen moeten worden overgelaten aan de volksvertegenwoordiging en het externe evaluatie-onderzoek aan een groep van onafhankelijke onderzoekers. Uiteraard zal tijdens de beleidsevaluatie gebruik gemaakt kunnen worden van de gegevens die het externe evaluatie-onderzoek heeft opgeleverd. Voorts kunnen de verkregen gegevens ook dienstig zijn voor andere functies (interne evaluatie van gebieden/scholen, ondersteuning). De roep om dienstbaarheid ten bate van deze functies zal echter nooit mogen resulteren in een aantasting van het plan van onafhankelijke, externe, evaluatie. Wat de uitvoering betreft merken zij op dat de vaststelling van het evaluatieplan dient te geschieden door de Tweede Kamer en dat de uitvoering in handen dient te worden gegeven van een coördinator die belast wordt met de detaillering van de onderzoeksplannen, de organisatie van het onderzoek en de verslaggeving. Het Centraal Instituut voor Toetsontwikkeling (CITO) wordt als meest geschikte locatie gezien voor de gegevensverzameling.

Het voorstel van Hofstee e.a. (1984) benadrukt, naar verwachting, het belang van onderzoek naar de vraag of de door het beleid beoogde *effecten*, te weten het opheffen of verminderen van onderwijsachterstanden, optreden. Omdat elke school met een bepaald percentage achterstandsleerlingen kan rekenen op extra faciliteiten, is een experimenteel design in strikte zin (experimentele groep vergelijken met een controlegroep) niet mogelijk. Wel mogelijk zou zijn een quasi-experimenteel design met een post hoc analyse van groepen scholen die vergelijkbaar zijn op sommige punten maar onvergelijkbaar op andere (die zeer relevant zijn voor de vraagstelling). De onderzoekers zijn echter van mening dat deze weg niet moet worden ingeslagen omdat zorgvuldige monitoring door middel van peiling van ontwikkeling van achterstanden meer tegemoet komt aan de informatiewensen van diverse actoren (parlement, departement, regering, publieke opinie) dan de informatie verkregen door een quasi-experimenteel design.

Gekozen wordt derhalve voor een zeer grootschalig cohortonderzoek waarbij leerlingen uit de groepen 2, 5 en 8 van de basisschool om de drie jaar worden getest op hun prestaties in rekenen en taal. Tevens wordt telkenmale een intelligentietest afgenomen. Door de testafname om de drie jaar te herhalen kunnen scholen ook met zichzelf in de tijd vergeleken worden. Het feit dat dezelfde leerlingen om de drie jaar worden hertest, heeft ook tot gevolg dat de leerwinst van leerlingen nauwkeurig in kaart kan worden gebracht. Een dergelijke opzet staat dus momentane vergelijkingen toe *en* vergelijkingen in de tijd (bijvoorbeeld: gebiedsscholen met niet-gebiedsscholen; scholen met veel versus scholen met weinig faciliteiten; leerjaar 4 op tijdstip Y vergeleken met

leerjaar 4 op tijdstip X; enz.)

Voorgesteld wordt om een steekproef te trekken van 1 op de 20 basisscholen. Aangezien op deze wijze het aantal scholen met veel achterstandsleerlingen in de steekproef beperkt zal zijn, wordt aanbevolen een substantieel extra aantal scholen van dit type aan dit cohort toe te voegen. De steekproef zou dan bestaan uit een representatieve selectie van scholen casu quo leerlingen (de referentiegroep), waaronder zich ook tal van gefaciliteerde scholen bevinden, aangevuld met scholen met veel leerlingen uit achterstandssituaties. Teneinde ook de effecten op langere termijn te kunnen traceren dienen ook scholen voor voortgezet onderwijs waar de leerlingen naar doorstromen, in de steekproef te worden opgenomen.

Effecten zouden moeten worden waargenomen in termen van (minder) verwijzing naar het speciaal onderwijs, (minder) doubleren, (hogere) toetsprestaties in de vakken taal en rekenen in de basisschool, (hoger) advies voortgezet onderwijs, (vaker) keuze voor hogere vormen van voortgezet onderwijs, (minder) doubleren in het voortgezet onderwijs, (geringere) afstroom, (hogere) opstroom en (hogere) toetsprestaties in de vakken Nederlands, wiskunde en Engels in het voortgezet onderwijs voor de leerlingen uit achterstandssituaties. Tevens wordt in de afname van intelligentietests voorzien. Deze kunnen als pure effectmaat dienst doen (heeft het beleid invloed op de intelligentie-ontwikkeling?) maar ook als referentiepunt: in hoeverre presteren leerlingen van een school beneden of boven hun intelligentieniveau ten gevolge van de door de school gekozen aanpak?

Teneinde te kunnen nagaan hoe beleidsmaatregelen door de overheid (extra faciliteiten) worden vertaald in maatregelen op school- en gebiedsniveau en hoe die zich verhouden tot de effecten, wordt implementatie-onderzoek voorgesteld. Variaties in implementatie op schoolniveau worden gerelateerd aan door de school behaalde resultaten. Nagegaan kan dan worden of sommige activiteiten op school- of gebiedsniveau effectiever zijn dan andere. Gegevens zullen derhalve moeten worden verzameld op landelijk-, gebieds- en schoolniveau. Ten slotte wordt in aanvulling op het grootschalige cohortonderzoek een aantal deelstudies voorgesteld onder meer betrekking hebbend op de zogenaamde 'brede' cognitieve- en sociale vaardigheden.

Door het Ministerie van Onderwijs en Wetenschappen wordt aan een drietal andere deskundigen een reactie gevraagd op het door Hofstee e.a. geleverde voorstel. Beem, Pijl en Scheerens (1986) maken enige kritische kanttekeningen bij het plan voor het volgen van alle cohortleerlingen, inclusief verwezenen en doubleurs (kosten!), en bij het gekozen interval van drie jaar. Fundamenteler is hun kritiek echter waar zij het ontbreken van adequate controlegroepen aan de orde stellen. Zij pleiten toch voor een (quasi-)experimenteel design. Hofstee (1986) stelt in een reactie dat het omvormen van de referentiegroep tot een (niet-equivalente, quasi-)controlegroep leidt tot een ernstige verzwakking van het evaluatie-ontwerp. De belangrijkste functie van de evaluatie voor de politieke en publieke discussie, namelijk het in de tijd peilen van onderwijsachterstanden, opgevat als discrepantie in gemiddelden tussen OVB leerlingen (leerlingen met een gewicht hoger dan 1.00) en de landelijke referentiegroep, zou daardoor opgeofferd worden (zie ook: Van der Werf & Tesser, 1990). In het definitieve plan is de argumentatie van Hofstee e.a. met betrekking tot het peilen in de tijd geaccepteerd en is de kritiek van Beem e.a. omtrent de cohortstructuur serieus genomen.

In dezelfde periode vraagt het Ministerie aan Doddema-Winsemius en Hofstee (1985) een nadere uitwerking van de effectmaten waar het gaat om leerprestaties, intelligentie en effecten in het sociaal-emotionele domein. Met betrekking tot de laatste stellen de onderzoekers dat 'sociale redzaamheid' hier het ankerbegrip is en dat het meten van houdingen binnen dit domein niet wenselijk is vanwege het probleem van indoctrinatie en vanwege de 'voorwendbaarheid' van houdingen.

In de voorbereidingsperiode bleken de voorgestelde intelligentietests in het overlegcircuit bij sommige actoren op grote weerstand te stuiten. Inderdaad bleek uit een haalbaarheidsonderzoek (Hoeben & Schaveling, 1986) dat veel schoolhoofden tegen afname van intelligentietests waren uit vrees voor stigmatisering van groepen leerlingen. Zoals uit het nog te beschrijven uitvoeringsplan zal blijken, zijn na

ampele discussies over voor- en nadelen de intelligentietests toch in het evaluatieplan opgenomen.

Eveneens in deze periode is aan onderzoekers van het Instituut voor Toegepaste Sociale Wetenschappen (ITS) gevraagd een gedetailleerder voorstel omtrent het implementatiedeel van het evaluatie-onderzoek, te weten: onderzoek naar het gebruik van de faciliteiten op school- en gebiedsniveau, uit te werken (Jungbluth, Buis, Driessen & Mens, 1985). De eerder genoemde onderzoekers Beem e.a. hebben in hun reactie (1986) ook dit voorstel kritisch onder de loep genomen en alternatieven aangedragen. In een poging de verschillende opties te overbruggen is Jungbluth en Meijnen (1986) gevraagd een resumerend advies uit te brengen. De aanbevelingen uit dit advies zijn in 1986 door de latere coördinator omgezet in een uitvoeringsplan (Kloprogge, 1986).

Inmiddels werd ook gewerkt aan de organisatorische vormgeving. De coördinatie wordt in handen gelegd van het Instituut voor Onderzoek van het Onderwijs (SVO). Een coördinator wordt benoemd (Kloprogge) die de onderzoeksplannen en begrotingen zal opstellen en dient zorg te dragen voor de contacten met de omgeving van het programma. Behalve het opstellen van meerjaren- en jaarlijkse plannen behoort ook het opstellen van een jaaroverzicht, een samenvatting van en een reflectie op de onderzoeksresultaten tot zijn taak. Voorts wordt er een Projectgroep Evaluatie OVB bestaande uit onafhankelijke deskundigen, ambtenaren van betrokken departementen en vertegenwoordigers van CITO en SVO door het Ministerie van O & W geïnstalleerd (voorzitter: Lagerweij), die als intermediair fungeert tussen onderzoekers en beleidsmakers. Ten slotte wordt een externe SVO commissie in het leven geroepen bestaande uit vier onafhankelijke deskundigen (voorzitter: Leune), die alle onderzoeksvoorstellen dient te beoordelen en worden twee onderzoeksinstituten geselecteerd (RION en ITS) die op grond van hun expertise het cohortonderzoek zullen gaan uitvoeren.

3 De voorbereiding geëvalueerd

In het voorafgaande is de feitelijke gang van zaken beschreven met betrekking tot de voorbereiding van het Landelijk Evaluatie Onderzoek naar het Onderwijsvoorrangsbeleid (LEO-OVB). Terwille van de evaluatie van de voorbereiding is het wenselijk te refereren aan de literatuur over dat type evaluatie dat het meest verwant is met, of geëigend is voor, het in gang gezette beleid.

Het Onderwijsvoorrangsbeleid is een landelijk beleidsprogramma en het beoordelen van de opbrengst daarvan zou daarom het best omschreven kunnen worden met de term Beleidsevaluatie. Hoogerwerf (1992) omschrijft dat laatste als: "Beleidsevaluatie is het beoordelen van de waargenomen inhoud, processen of effecten van een beleid aan de hand van bepaalde criteria" (p. 9).

Een dergelijke evaluatie kan zowel ex ante als ex post plaatsvinden. In het eerste geval gaat het om een beoordeling van het voorgenomen beleid, ook wel analytische evaluatie genoemd. Rationaliteit en legitimiteit van doelstellingen alsmede de doel-middelen structuur van het beleidsprogramma worden aan beoordeling onderworpen in het licht van de beschikbare wetenschappelijke kennis.

Een expliciete ex ante evaluatie heeft niet plaatsgevonden, hoewel die in het uitvoeringsplan wel was opgenomen. Naar de reden daarvan kan men gissen. Wellicht wordt de behoefte daaraan door beleidsfunctionarissen het sterkst gevoeld als geheel nieuw beleid wordt overwogen. Dat was hier niet het geval: het OVB kan als een gewijzigde vorm van het Stimuleringsbeleid (1975-1984) worden beschouwd. Wellicht wordt ook de kritische functie die een analytische evaluatie kan hebben door beleidsfunctionarissen niet altijd op prijs gesteld.

Ex post evaluaties van beleidsprogramma's zoals het Onderwijsvoorrangsbeleid (een casus, met wellicht een vergelijkingsmogelijkheid in de tijd: het Stimuleringsbeleid) zijn er in vele soorten maar de volgende vragen zijn vrijwel altijd aan de orde (Hoogerwerf, 1992, p. 55, citaat):

"a. Welke zijn de doeleinden en instrumenten van het beleid? En welke evaluatiecrite-

ria kunnen aan de beleidsdoeleinden worden ontleend?

b. In hoeverre zijn de doeleinden van het beleid bereikt? En welke niet beoogde effecten treden op?

c. In hoeverre valt de mate van doelbereiking te verklaren uit het beleid als geheel en de gekozen instrumenten in het bijzonder?

d. Hoe verhouden zich de kosten (negatieve effecten) en de baten (positieve effecten) tot elkaar?

e. Uit welke andere factoren dan de beleidsinhoud vallen de mate van doelbereiking, effectiviteit en efficiëntie te verklaren? (...) Bij die andere factoren denkt men vooral aan het beleidsproces (met name de beleidsvoering), de beleidsorganisatie, het beleidsveld en de kenmerken van de afzonderlijke factoren, waaronder hun doelgerichtheid, informatie en macht."

Wanneer we het uitvoeringsplan (Kloprogge, 1986) als afronding van de voorbereiding van het evaluatieplan beschouwen en de dan voorliggende opzet confronteren met de door Hoogerwerf als gangbaar omschreven vragen inzake beleidsevaluatie, dan kunnen we het volgende vaststellen. Met betrekking tot vraag b wordt in het uitvoeringsplan wel aandacht geschonken aan het vraagstuk van de niet beoogde effecten maar worden daar geen expliciete hypothesen over geformuleerd. Dataverzameling specifiek hierop gericht is dan ook niet gepland. Voorts komen de vragen c en d niet of nauwelijks aan de orde. Er wordt met andere woorden in het evaluatieplan niet of nauwelijks voorzien in onderzoek anders dan naar de inzet van de middelen en de effecten op de beoogde doelen (leerprestaties, schoolcarrières). Zoals reeds aangegeven gaven Hofstee e.a. zelf al aan dat hun plan *geen* beleidsevaluatie inhoudt, beleidsevaluatie omvat in hun ogen meer en zou overgelaten moeten worden aan de volksvertegenwoordiging. De laatste zinsnede is meerduidig. Dat het primaat in de besluitvorming, en dus de finale beoordeling, over nationale beleidsvragen bij de politiek ligt en niet bij de wetenschap staat niet ter discussie. Men zou de zinsnede echter ook zo op kunnen vatten dat wat wetenschappelijk omtrent het OVB te evalueren valt wel zo ongeveer is opgeschreven in het door hen gepresenteerde plan. Dat laatste is echter niet het geval. Er valt wel degelijk wetenschappelijk verantwoord onderzoek te verrichten naar de inhoud, het krachtenveld, de organisatie en het proces van het beleidsprogramma (zie o.a. Karstanje, 1988; Verhoeff, 1992). En ook onderzoek naar niet beoogde effecten behoort zeker tot de mogelijkheden.

Wanneer we de voorbereiding evalueren, dienen we nog bij enkele zeer relevante methodische punten stil te staan. Als eerste is dat de discussie tussen Hofstee e.a. en Beem e.a. omtrent de wenselijkheid van adequate controlegroepen. In de evaluatieliteratuur is dit probleem bekend in termen van de spanning tussen externe en interne validiteit (Cronbach, 1983; Cook & Campbell, 1979). Anders gezegd: kiest men voor een opzet waarbij het funtioneren van het programma in de context zo goed mogelijk wordt gedocumenteerd (externe validiteit) dan wel voor een ontwerp waarbij men het oorzakelijk verband tussen een activiteit en het effect binnen *dit* experiment wil aantonen (interne validiteit). De positie van Beem e.a. is te omschrijven als, in navolging van Cook en Campbell, het laten prevaleren van de interne validiteit terwijl Hofstee e.a. overeenkomstig Cronbach, de externe validiteit van de te vergaren gegevens benadrukken door te wijzen op de functie die de gegevens, verkregen door hun onderzoeksopzet, in de politieke en publieke discussie zullen vervullen.

Het debat over de functie van controlegroepen in programma-evaluatie heeft de laatste tijd iets van zijn controverse verloren. In een recent overzichtsartikel omtrent de rol van quasi-experimenten in de evaluatie van beleidsprogramma's komt Cook (1991) tot de volgende conclusie: "In 1986, Campbell's work on quasi-experiments was still the most widely acknowledged method source among evaluators. However, our impression is that by then it had lost the hegemony it had once enjoyed. It is now among the methods to be used when descriptive causal questions can be justified as the centerpiece of an evaluation design. But it is definitely not the sole method of choice" (p. 141). Naar aanleiding van een presentatie van een paper van Kloprogge over LEO-OVB in Washington in 1991 constateerde Campbell (!) dan ook dat naar zijn oordeel de gekozen opzet voor de Nederlandse situatie de meest adequate was (mondelinge mededeling).

Een tweede relevant punt is het ontbreken van criteria ('standaarden') in het evaluatieplan. Nergens wordt aangegeven welke winst in leerprestaties of schoolcarrières leerlingen uit achterstandssituaties moeten boeken om van een succesvol beleid te kunnen spreken. Het onderwerp komt noch in het ontwerp van Hofstee e.a. noch in het uitvoeringsplan ter sprake. Uit de zinsnede dat het peilen van onderwijsachterstanden in de tijd de belangrijkste functie is van de externe evaluatie, zou men kunnen opmaken dat de auteurs kennelijk de effecten van dit specifieke beleid niet denken te kunnen schatten dan wel te kunnen isoleren en het derhalve ook geen zin heeft om een nauwkeurig omschreven standaard te formuleren.

Tot slot valt te concluderen dat de voorbereiding van het evaluatie-onderzoek vergelijkenderwijs, zeer grondig is geweest. Contra-expertises zijn verricht, uitwerkingsopdrachten zijn verleend, resumerende adviezen zijn gevraagd en er is een infrastructuur gecreëerd die de onafhankelijkheid en de professionele uitvoering van de evaluatie zo goed als mogelijk garanderen.

4 De uitvoering: 1986-1992

Zoals al eerder is aangegeven, is het resultaat van de voorbereidingsfase neergelegd in een uitvoeringsplan (Kloprogge, 1986). Voorts is in een themanummer van dit tijdschrift getiteld: *Evolutie in Onderwijsevaluaties* door Van der Werf en Tesser (1990) al redelijk gedetailleerd verslag gedaan over de opzet en uitvoering tot halverwege 1989. Ook verschijnt er elk jaar een door de coördinator geschreven samenvattende rapportage (Kloprogge, 1988, 1989, 1990, 1991, 1992). Wij zullen ons daarom beperken tot een bespreking op hoofdlijnen die zich leent voor een kritische beschouwing.

Eveneens al in het voorgaande aangeduid kan LEO-OVB opgedeeld worden in twee hoofdstromen van onderzoek: (a) het cohortonderzoek en (b) het flankerend onderzoek. Voor het totaal is per jaar circa 1,5 miljoen gulden beschikbaar.

a. *Het cohortonderzoek*

Om de twee jaar worden bij de leerlingen uit de groepen 4, 6 en 8 van circa 700 basisscholen toetsen voor taal, rekenen en een intelligentietest afgenomen. Hiermee is gestart in het najaar van 1988. Aan het oorspronkelijke voornemen om in deze cyclus ook een test voor sociale redzaamheid op te nemen, kon geen uitvoering worden gegeven omdat de instrumentconstructie teveel problemen opleverde. Per leerjaar betreft het ongeveer 12000 leerlingen. Aansluitend op de dataverzameling bij de leerlingen worden door middel van vragenlijsten gegevens verzameld over de school als organisatie en over het curriculum. Met betrekking tot het laatste wordt een onderscheid gemaakt naar kenmerken die uit de literatuur over schooleffectiviteit belangrijk zijn gebleken en kenmerken die vanuit de OVB-literatuur van belang werden geacht. Voorts werd de ouders gevraagd een vragenlijst in te vullen omtrent relevante gezinskenmerken (opleiding, beroep, land van herkomst enz.).

Het onderzoek in het voortgezet onderwijs is gestart een jaar nadat de eerste golf leerlingen uit groep 8 de basisschool had verlaten, dus in 1989. Dit betrof een random steekproef van 50% van deze leerlingen. Zij werden weer opgespoord en de overgang naar het voortgezet onderwijs evenals hun prestaties (cijfers) werden in kaart gebracht. Een jaar later (1990) is hun onderwijspositie opnieuw geïnventariseerd. Inmiddels zijn de leerlingen van de groepen 4, 6 en 8 van de basisscholen zowel in 1990 als in 1992 weer getest en getoetst. Van een generatie leerlingen zijn dus nu de gegevens over de gehele basisschoolperiode verzameld. Voorts zijn in 1990 en 1992 opnieuw cohorten gestart in groep 4. Onderzoek naar verwezenen naar het Speciaal Onderwijs is eveneens gedaan binnen het cohort 1988. Evenals onderzoek naar schoolverlaters binnen het cohort in het voortgezet onderwijs.

b. *Het flankerend onderzoek*

In de loop der tijd is uit het LEO-OVB budget ook veel onderzoek verricht naar specifieke problemen het OVB betreffende. Elk jaar legt de coördinator van LEO-OVB een lijst met onderwerpen voor aan de Projectgroep Evaluatie OVB die naar zijn mening voor onderzoek in aanmerking komen. Voorts reserveert hij een

budget waarmee beloftevolle, plaatselijke initiatieven geëvalueerd kunnen worden. Ook deze initiatieven worden voorgelegd aan de Projectgroep. In totaal zijn binnen de categorie flankerend onderzoek 15 projecten uitgevoerd of in uitvoering. Deze worden niet alleen bij het RION of het ITS aanbesteed maar ook bij andere onderzoeksinstituten of vakgroepen van universiteiten.

Voorbeelden van flankerend onderzoek zijn: het functioneren van OVB-gebieden, de relatie tussen OVB en de basiseducatie, de aspiraties en percepties van buitenlandse ouders en hun kinderen, de evaluatie van specifieke projecten voor buitenlandse meisjes, zorgverbreding technisch lezen, een onderzoek naar de ondersteuning.

De onderzoeksvoorstellen voor zowel het cohortonderzoek als het flankerend onderzoek worden, zoals gepland, voorgelegd aan een externe beoordelingscommissie van SVO en aan de Projectgroep. De eerste beoordeelt met name de researchtechnische aspecten van het voorstel en de laatste de uitwerking van de probleemstelling. Jaarlijks worden de samenvattende rapportages van de coördinator aangeboden aan de betrokken ministeries en aan het parlement. Voorts worden de resultaten van de test- en toetsafnames teruggerapporteerd aan de scholen.

De uitvoering van LEO-OVB loopt, in ruime zin, vanaf 1986 en in engere zin vanaf 1988 als we de start van het cohortonderzoek als uitgangspunt nemen, dus respectievelijk 6 en 4 jaar. Lang genoeg om er in evaluatief opzicht enige kanttekenigen bij te plaatsen.

5 Evaluatieve kanttekeningen bij de uitvoering 1986-1992

De omgeving van een vrij groot evaluatieprogramma als van het OVB, is in principe vrij uitgebreid. Met bijna alle instellingen en organisaties, die werkzaam zijn op het onderwerp van onderzoek, zijn er wel relaties. Voor het functioneren van het evaluatieprogramma zijn er echter drie categorieen uitermate relevant:

a de opdrachtgever, casu quo het Ministerie van Onderwijs en Wetenschappen;
b de wetenschap;
c de geëvalueerden.

In veel literatuur over evaluatie worden vooral de opdrachtgever en de wetenschap als relevante actoren gezien (JCSEE, 1981; Chelimsky, 1987). In de Nederlandse situatie met zijn unieke verzorgingsstructuur voor het onderwijs is het echter dringend aan te bevelen ook degenen die worden geëvalueerd tot het krachtenveld rond de evaluatie te rekenen.

Deze drie elementen in de omgeving van een evaluatieprogramma hebben enkele karakteristieken die voor de evaluatie van belang zijn:

- elk van de drie is geen eenheid op zich; er is een behoorlijke mate van heterogeniteit;
- elk van de drie heeft unieke inzichten, denkbeelden en belangen;
- er is sprake van relaties tussen deze drie omgevingsfactoren.

De heterogeniteit van de elementen in de omgeving van de evaluatie is wellicht het moeilijkst te hanteren vanuit een evaluatieprogramma. De evaluatoren moeten er rekening mee houden dat er voortdurend verschuivingen optreden zowel bij de opdrachtgever als in de wetenschap en in het veld van onderzoek.

Wat betreft de opdrachtgever is voor de evaluatie van het OVB steeds van belang geweest in hoeverre verschillende onderdelen van het Ministerie of individuele ambtenaren vonden dat de evaluatie ook hun werkterrein raakte. In de wetenschap zijn er uiteenlopende disciplines en wetenschapsopvattingen, die in de loop der tijd krachtiger of zwakker worden. Tijdens de uitvoering van de evaluatie kwam het onderzoek naar de effectiviteit van scholen en instructie sterk opzetten en raakte het kwalitatieve onderzoek deels uit de gratie. In het object van onderzoek treedt er ook voortdurend een wisseling op van machtsposities en aanspreekpunten. Toen het Ministerie van WVC zich uit het OVB tijdelijk terugtrok leidde dat bijvoorbeeld tot een verzwakking van de positie van de welzijnsinstellingen ten opzichte van de scholen. Met de komst van het beleid van Sociale Vernieuwing en met een hernieuwde aandacht van WVC voor het OVB hebben er zich in het veld opnieuw wisseling van machtsposities voorgedaan.

De eigenheid aan kennis, inzichten en belangen bij de drie elementen in de omgeving is

vanuit de organisatie van het programma minder moeilijk te hanteren. Aangezien de evaluatie zelf een wetenschappelijke activiteit is, behoort het volgen van de ontwikkeling van kennis en inzichten uit de wetenschap tot de reguliere taak van de evaluatie-onderzoekers. Ambtenaren op een departement beschikken vaak over informatie over te verwachten beleidsontwikkelingen, waar de evaluatie op kan of moet anticiperen. Degenen die worden geevalueerd hebben veelal een gedetailleerd inzicht in de praktijk van het dagelijks werk, dat tot de evaluatie behoort. Dat kan zowel leiden tot nieuwe aandachtsvelden voor de evaluatie als bijdragen aan het bevestigen van de juistheid van de onderzoeksresultaten of het relativeren daarvan. In deze zin is de in de omgeving aanwezige informatie van groot belang om de kwaliteit van de evaluatie te verhogen.

Aan in de omgeving spelende belangen kan ook de evaluatie niet voorbij gaan. De meest geëigende oplossing daarvoor is bij de opzet van de evaluatie daarop te anticiperen zowel met betrekking tot de inhoud van het programma als met de structurering van de contacten tussen het programma en de omgeving. In het voorgaande is deze structuur al uiteengezet. Dat de evaluatie gedurende een aantal jaren vrij continu en ongestoord kon worden uitgevoerd, kan als een indicatie gelden voor de kwaliteit van de structurering van de omgeving van het programma.

Een goede structurering is echter geen panacee voor alle problemen. In een betrekkelijk kleinschalige omgeving als de Nederlandse, waar alle belangrijke actoren elkaar kennen en in allerlei circuits elkaar ontmoeten, zijn de drie elementen in de omgeving duidelijk vervlochten. Dit houdt in dat een departement als opdrachtgever zeer terughoudend zal zijn om evaluatieve activiteiten te financieren, die in wetenschappelijke kring omstreden zijn of waar in het veld van onderzoek grote bezwaren tegen bestaan. Omgekeerd is het moeilijk voor de onderzochten om bezwaar te maken tegen evaluatie-onderzoek, dat wordt gesteund vanuit het departement en waarvan wetenschappers het belang onderstrepen. De vervlechting die er bestaat tussen de actoren in de omgeving kan dus zowel de positie van de evaluatoren verzwakken als versterken. Het is in ieder geval van belang met deze vervlechting vanuit managementoogpunt rekening te houden. Dit houdt onder meer in dat er heel snel moet worden gereageerd als er ergens tegen een programma-onderdeel verzet rijst. Bij het begin van het cohortonderzoek bleken er bijvoorbeeld in enkele scholen in een gemeente bezwaren te leven tegen het afnemen van toetsen, hoewel ze hadden ingestemd met de deelname aan het programma. Deze scholen zochten medestanders in de gemeente en daarbuiten, waaronder wetenschappers. Dergelijke ontwikkelingen kunnen snel de omvang aannemen van een onoplosbaar probleem. Door in een heel vroeg stadium contact op te nemen met de betrokken scholen, uitleg te geven over de feitelijke gang van zaken en voorzieningen te treffen op die punten waar de argumenten van de scholen ter zake bleken, konden de problemen worden opgelost.

Voor het zodanig structureren van de omgeving van het programma dat het een van tevoren gegarandeerd succes wordt, bieden wetenschappelijke kennis en onderzoekservaring weinig aanknopingspunten. Belangrijk is echter dat de voorgenomen onderzoeksactiviteiten als relevant en kwalitatief hoogwaardig worden beschouwd. Hierdoor wordt een basis gelegd op grond waarvan de interactie met de omgeving kan plaatshebben. Het blijft echter nodig in het onderzoek gemaakte keuzen expliciet en veelvuldig te verantwoorden, en het is van groot belang ervoor te zorgen dat voor dergelijke activiteiten tijd en geld beschikbaar is.

Na deze meer algemene opmerkingen zal in het kort worden ingegaan op de relatie van het programma met elk der drie categorieën.

a. *De opdrachtgever*

In de praktijk blijkt het Ministerie de onderzoeksvoorstellen van de Projectgroep ongewijzigd over te nemen, waarbij het belang van bepaalde onderdelen soms expliciet wordt onderstreept. Commentaar op de kwaliteit van de onderzoeksresultaten wordt van de zijde van het Ministerie, de verantwoordelijke bewindslieden en het parlement niet gehoord. De resultaten op zich hebben aanleiding gegeven zowel tot commentaar van parlementariërs als van beleidsvoornemens door staatssecretaris Wallage. De laatste heeft naar aanleiding van de bevinding dat meer dan 45% van de leerlingen tot de doelgroepen van het beleid behoort, beslo-

ten tot aanscherping van de criteria. Voorts heeft hij op grond van de peiling van de achterstand van groepen allochtone leerlingen door LEO-OVB een commissie onder voorzitterschap van Van Kemenade gevraagd alternatieven, dan wel wijzigingen, ten opzichte van het huidige beleid te formuleren. De genoemde commissie heeft bij de analyse van het probleem uitvoerig gebruik gemaakt van de gegevens van het evaluatieprogramma (Ministerie van Onderwijs en Wetenschappen, 1992a,b). Voorts wordt in de Toelichting op de Wet op het Onderwijssvoorrangsbeleid, die momenteel (januari 1993) nog in behandeling is bij de Eerste Kamer, veelvuldig verwezen naar de resultaten verkregen via LEO-OVB.

b. *De wetenschap*

De onderzoeksvoorstellen worden, zoals gezegd, voorgelegd aan een commissie van onafhankelijke deskundigen. Het komt maar sporadisch voor dat een eerste voorstel genade vindt in de ogen van de commissie. Voorstellen worden herschreven met in sommige gevallen alsnog een negatief resultaat. Eindverslagen worden gepubliceerd in een afzonderlijke reeks na eerst te zijn becommentarieerd door de Projectgroep. Voorts verschijnen er bijdragen van de onderzoekers in wetenschappelijke tijdschriften, zowel nationaal als, sporadisch, internationaal en worden papers gepresenteerd op wetenschappelijke congressen. Gelet op de kwaliteit, omvang en longitudinale structuur van de datasets van de cohorten moet echter worden opgemerkt dat de wetenschappelijke potentie daarvan bij lange na niet wordt benut. Er gaat zoveel menskracht zitten in het verzamelen van de data en het globaal analyseren daarvan ten behoeve het beleid, dat tal van wetenschappelijke vragen die ook middels deze data geanalyseerd kunnen worden, niet aan bod komen (stabiliteit van prestaties van scholen, stabiliteit van prestaties van leerlingen, kwaliteit versus egaliserend vermogen van scholen, detailanalyses van specifieke groepen leerlingen enz.). Gelet op de kwaliteit en structuur van de dataset en gegeven de aanwezigheid van hoogwaardige analyse-expertise in Nederland zou ook in buitenlandse tijdschriften veel meer over het evaluatieprogramma gepubliceerd kunnen worden dan tot nu toe het geval is.

c. *De geëvalueerden*

Zoals gezegd, bestonden er bij de start van het onderzoek op veel scholen bezwaren tegen de afname van tests en toetsen. Na de eerste terugrapportage aan de scholen van zowel de toetsgegevens (gespecificeerd per school) als de samenvattende rapportages, ontstond echter een niet verwachte appreciatie in het veld. Tal van uitnodigingen bereikten de onderzoekers om de resultaten te komen toelichten dan wel om bijvoorbeeld analyses op gebiedsniveau te verrichten. Al spoedig bleek de vraag dusdanig aan te zwellen dat de inrichting van een 'transferpunt' noodzakelijk werd. Met instemming van het Ministerie van O en W is dat ook gebeurd en is aanvullende formatie ter beschikking gesteld. Spoedig daarna vroegen ook bevoegde gezagen van scholen buiten de steekproef of het toetsprogramma ook bij hun scholen, tegen betaling, kon worden afgenomen. Er was al spoedig sprake van olievlekwerking: in grote steden als Rotterdam en Den Haag worden momenteel alle scholen uit onderwijsvoorrangsgebieden onderworpen aan het toetsprogramma.

De bevoegde gezagen gebruiken deze gegevens voor interne evaluatiedoeleinden. Wat aanvankelijk een hindernis leek voor het evaluatieprogramma, verkeerde in haar grootste aantrekkingskracht: het rapporteren over effecten. Bijgevolg is ook de steekproefmortaliteit bijzonder gering en de medewerking van de scholen in de steekproef zeer groot. Een dergelijke wending is uiteraard niet te voorzien en zal in dit geval alles te maken hebben met de toenemende aandacht voor accountability in het onderwijs.

Resumerend: de relaties van het evaluatieprogramma met de belangrijkste actoren zijn zonder meer goed te noemen.

Ten slotte zullen we kort ingaan op een vraag die in veel handboeken over evaluatie aan de orde komt, namelijk die naar het relatieve belang van kwaliteit versus tijdigheid. Evaluaties die voor een opdrachtgever worden uitgevoerd, dienen meestal om beslissingen te nemen over voortzetting of bijstelling van door de opdrachtgever uitgevoerde of gefinancierde activiteiten. Het tijdig opleveren van rapportages behoort dan ook tot de standaarden die voor

evaluaties gelden. Chelimsky (1987) trekt uit haar ervaringen de les dat een adequaat design, waarbij de gevraagde resultaten tijdig kunnen worden opgeleverd, vaak de voorkeur verdient boven een optimaal design. Cronbach (1983) lijkt de factor tijd meer te zien als een van de vele randvoorwaarden die in een evaluatieprogramma zo goed mogelijk moeten worden verzoend. Scheerens (1983) vindt het wel van belang dat voornemens nauwgezet en volgens de tijdsplanning worden nagekomen als onderdeel van het streven dat projectmanagement van beleidsgerichte onderwijsresearch in het algemeen moet worden verbeterd.

In de evaluatie van het OVB heeft het werken volgens een strak tijdsschema een hoge prioriteit gekregen. Er hebben zich echter diverse situaties voorgedaan waarbij een afweging moest worden gemaakt tussen kwaliteit en tijdigheid. In het algemeen luidt de conclusie hier dat het is aan te bevelen de hoogste prioriteit toe te kennen aan kwaliteitsaspecten. Tijdigheid is rekbaarder dan kwaliteit. Harde termijnen waarop activiteiten worden gepland blijken achteraf toch vaak minder hard te zijn. Politieke besluitvorming kost de nodige tijd en verloopt in diverse fasen. Het al te strak vasthouden aan termijnen is daarom vaak niet zinvol. Wanneer gekozen wordt voor minder optimale designs of analyses, levert dit meestal moeilijk herstelbare problemen op. Tijdigheid blijft uiteraard belangrijk. Aantasting van de kwaliteit ter wille van de tijdsplanning lijkt echter zelden te lonen bij evaluatie van een beleidsprogramma dat zich over vele jaren uitstrekt.

6 Conclusies

Samenvattend kan worden gesteld dat het van belang is om bij de opzet van de evaluatie van een beleidsprogramma nauwkeurig na te gaan welke relevante actoren zich in de omgeving van het programma bevinden en welke eisen het omgaan met de omgeving aan het programma stelt. Het zorgvuldig opzetten van het onderzoek en het structureren van de omgeving van het programma kost tijd en geld maar ze kunnen de kwaliteit van de evaluatie sterk ten goede komen. Het zorgen voor hoogwaardige kwaliteit van het onderzoek blijft de eerste prioriteit van een evaluatie van een beleidsprogramma en is ook voorwaarde om binnen een complexe omgeving te kunnen werken. Het werken binnen gestelde termijnen is voor een evaluatieprogramma eveneens van belang maar lijkt zelden voorrang te verdienen boven het kwaliteitsaspect. In het bijzonder geldt dat voor programma's als het OVB waar niet zozeer voor een eenmalige beleidsinzet wordt geopteerd maar voor beleid dat, weliswaar in meerdere vormen, decennia bestrijkt. Vanuit deze optiek is programma-evaluatie niet in eerste instantie de evaluatie van een in de tijd goed gedefinieerd beleidsproces maar de registratie van ontwikkelingen in de tijd ('rolling reform'). Een zorgvuldige monitoring van deze ontwikkelingen, in dit geval van prestaties en schoolcarrières van groepen leerlingen, gecombineerd met zo goed mogelijke analyses richting oorzaken, is dan te prefereren boven reeksen dieptestudies dan wel experimentele designs die een precieze tijdsbegrenzing vereisen.

Tot slot. Ook binnen de categorie 'flankerend onderzoek' treffen we geen onderzoek aan dat het beleidsproces in 'hogere sferen' als onderwerp heeft. Het feit dat de Onderwijsvoorrangswet na acht jaar nog steeds niet is aangenomen en de gevolgen daarvan, de wijzigingen in doelstellingen van het OVB in de loop der tijd, de werkgelegenheidsconsequenties van de maatregelen enz. zijn geen voorwerp van onderzoek geweest. In die zin is ook het LEO-OVB geen volledige beleidsevaluatie in de zin die Hoogerwerf (1992) er aan toekent. Prioriteiten in deze richting van de kant van de opdrachtgever zijn niet te verwachten. We zullen het hier moeten hebben van onderzoek gefinancierd door NWO of door een vakgroep van een universiteit.

Literatuur

Beem, A. L., Pijl, Y. J., Scheerens, J. (1986). *Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma Onderwijsvoorrangsbeleid (OVB).* Den Haag: SVO.

Chelimsky, E. (1987). What have we learned about the politics of program evaluation? *Educational Evaluation and Policy Analysis, 9*(3), 199-213.

Cook, Th. D. (1991). Clarifying the warrant for generalized causal inferences in quasi-experimentation. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and Education: At Quarter Century.* (pp. 115-145). Chicago: The University of Chicago Press.

Cook, Th. D., & Campbell D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally.

Cronbach, L. J. (1983). *Designing Evaluations of Educational and Social Programs.* San Francisco: Jossey-Bass Publishers.

Doddema-Winsemius, H., & Hofstee, W. K. B. (1985). *Operationalisering van de effecten van het onderwijsvoorrangsbeleid (OVB).* Den Haag: SVO.

Hoeben, W. Th. J. G., & Schaveling, J. (1986). *Evaluatie van het onderwijsvoorrangsbeleid; Haalbaar?* Groningen: RION.

Hofstee, W. K. B., Koster, K. B., Meijnen, G. W. (1984). *Programmavoorstel voor evaluatieonderzoek naar de effecten van het onderwijsvoorrangsbeleid.* Groningen: Rijksuniversiteit.

Hofstee, W. K. B. (1986). *Reactie naar aanleiding van het eindrapport Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma OVB.* Groningen: vakgroep persoonlijkheidspsychologie (interne notitie).

Hoogerwerf, A. (1992). Naar een verbreding van de beleidsevaluatie. *Beleidswetenschap, 6*(1), 52-68.

JCSEE, (1981). *Standards for Evaluation of Educational Programs, Projects and Materials.* New York: McGraw-Hill.

Jungbluth, P., Buis, Th., Driessen, G., Mens, A. (1985). *De evaluatie van het onderwijsvoorrangsbeleid. Voorstellen voor onderzoek naar het gebruik van OVB-faciliteiten en naar de daarbij beoogde effecten.* Nijmegen: ITS.

Jungbluth, P., & Meijnen, G. W. (1986). *Resumerend advies over de vormgeving van de evaluatie van het onderwijsvoorrangsbeleid.* Nijmegen: ITS.

Karstanje, P. N. (1988). *Beleidsevaluatie bij controversiële onderwijsvernieuwing.* Amsterdam: SCO.

Kloprogge, J. (1986). *Uitvoeringsplan evaluatie onderwijsvoorrangsbeleid 1987-1990.* Den Haag: SVO.

Kloprogge, J. (1988). *En toen was er het OVB...? Notitie betreffende het onderwijsvoorrangsbeleid in 1987.* Den Haag: SVO.

Kloprogge, J. (1989). *Voorrang voor de gebieden? Notitie over het onderwijsvoorrangsbeleid in 1988.* Den Haag: SVO.

Kloprogge, J. (1990). *Naar beter onderwijsvoorrangsbeleid. Notitie over het onderwijsvoorrangsbeleid in 1989.* Den Haag: SVO.

Kloprogge, J. (1991). *Het gewicht gewogen. Notitie over het onderwijsvoorrangsbeleid in 1990.* Den Haag: SVO.

Kloprogge, J. (1992). *Een kompas voor de toekomst. Notitie over het onderwijsvoorrangsbeleid in 1991.* Den Haag: SVO.

Mens, A., & Calcar, C. van. (1981). *Onderwijsstimulering op weg.* Nijmegen/Amsterdam: ITS/SCO.

McLaughin, M. W., & Phillips, D. C. (Eds.). (1991). *Evaluation and Education: At Quarter Century. Ninetieth yearbook of the National Society for the Study of Education.* Chicago: The University of Chicago Press.

Ministerie van Onderwijs en Wetenschappen (1984). *Vaststelling van een wettelijk kader voor het voeren van een samenhangend onderwijsvoorrangsbeleid (onderwijsvoorrangswet).* Den Haag: Staatsuitgeverij.

Ministerie van Onderwijs en Wetenschappen (1985). *Onderwijsvoorrangsplan 1985-1989.* Den Haag: Staatsuitgeverij.

Ministerie van Onderwijs en Wetenschappen (1992a). *Ceders in de tuin. Naar een nieuwe opzet van het onderwijsbeleid voor allochtone leerlingen. Deel I.* Den Haag: Sdu.

Ministerie van Onderwijs en Wetenschappen (1992b). *Ceders in de tuin. Voorstudies. Deel II.* Den Haag: Sdu.

Scheerens, J. (1983). *Evaluatie-onderzoek en beleid.* Den Haag: SVO.

Verhoeff, C. C. (1992). *Netwerkgebonden beleidsprogramma's in het onderwijs. Een studie naar implementatie en doelrealisatie van veranderingen in scholenbestand en bekostiging.* De Lier: Academisch Boeken Centrum.

Werf, M. P. C. van der, & Tesser, P. (1990). De evaluatie van het onderwijsvoorrangsbeleid. *Pedagogische Studiën, 67,* 127-139.

Manuscript aanvaard 11-1-1993

Auteurs

G. W. Meijnen is hoogleraar onderwijskunde aan de Universiteit van Amsterdam en wetenschappelijk directeur van het SCO-Kohnstamm Instituut.

Adres: SCO-Kohnstamm Instituut, Grote Bickerstraat 72, 1013 KS, Amsterdam.

J. Kloprogge is coördinator van de landelijke evaluatie van het onderwijsvoorrangsbeleid bij SVO.

Adres: SVO, Sweelinckplein 14, 2517 GK, Den Haag.

Abstract

The Evaluation of the Educational Priority Policy in The Netherlands Evaluated

G. P. Meijnen & J. Kloprogge. Pedagogische Studiën, 1993, 70, 84-96.

This article presents the analysis of a national evaluation study regarding the Educational Priority Policy. Both, the preliminary phase and the implementation are described. Furthermore, the article focuses on those topics which are considered to be important in international literature, namely, internal versus external validity, design choices and interrelations between the research programme and the environment. It can be concluded that both design and the way the research programme is carried out, still meet recent views regarding external evaluation research. At the same time, it has been established that policy evaluation encompasses more than the present contents of the research programme.

De prestaties van de OVB-doelgroepen in 1988: stand van zaken na de eerste meting*

M. P. C. van der Werf, P. Tesser en L. Mulder

Samenvatting

Op basis van de gegevens van de eerste meting van de evaluatie van het onderwijsvoorrangsbeleid in 1988 worden verbanden geanalyseerd tussen de sociaal-etnische achtergrond, intelligentie en het thuisklimaat van leerlingen en hun taal- en rekenprestaties in groep 4, 6 en 8 van het basisonderwijs. Tevens worden verschillen tussen groepen scholen nagegaan. Zowel de taal- als de rekenprestaties van allochtone leerlingen blijven ver achter bij het landelijk gemiddelde. Voor de 1.25 leerlingen zijn de achterstanden veel kleiner. Er is een sterk verband met intelligentie en een zwak verband met de aanwezigheid van culturele hulpbronnen thuis. Na controle voor intelligentie wordt het effect van de weegfactor aanmerkelijk kleiner, na controle voor culturele hulpbronnen neemt dit effect slechts in geringe mate af.

Gebiedsscholen scoren duidelijk lager dan scholen met alleen extra basisformatie en deze scoren weer lager dan scholen zonder OVB-faciliteiten. Na controle voor de weegfactoren van de leerlingen, het opleidingsniveau van hun ouders en het herkomstland van de leerlingen is er alleen nog een effect van de gebiedsscholen voor de 1.25 leerlingen. Dit effect blijft ook bestaan na controle voor intelligentie. Uit andere analyses is gebleken dat dit effect vrijwel volledig kan worden toegeschreven aan het hoge percentage 1.90 leerlingen op de gebiedsscholen.

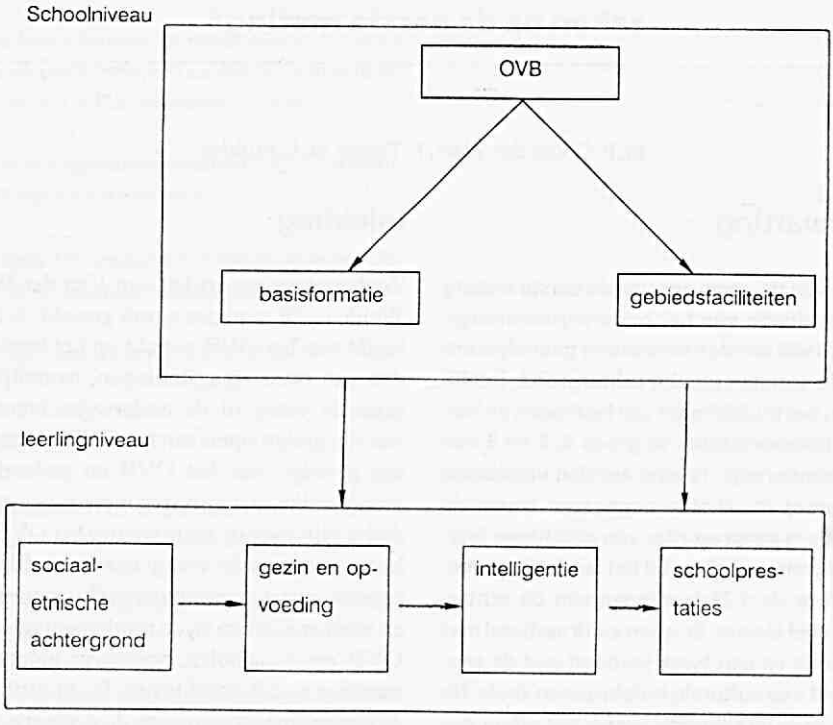
Inleiding

Zoals ook in het artikel van Van der Werf en Weide in dit nummer wordt gesteld, is de evaluatie van het OVB gericht op het beantwoorden van twee vraagstellingen, namelijk enerzijds de vraag of de onderwijsachterstanden van de doelgroepen van het OVB verminderen ten gevolge van het OVB en anderzijds de vraag welke factoren in concreto verantwoordelijk zijn voor de mate waarin het OVB effect heeft. De tweede vraag wordt hoofdzakelijk beantwoord door het onderzoek naar de school- en klaskenmerken en de implementatie van het OVB op de scholen, waarover elders in dit nummer wordt geschreven. In dit artikel staat de eerste vraagstelling van de evaluatie van het OVB centraal. Dat betekent dat de interesse vooral uitgaat naar de schoolprestaties van de doelgroepen van het OVB.

Verondersteld wordt dat, als het beleid werkt, de invloed van sociale en etnische achtergrondkenmerken op de schoolprestaties van de leerlingen op den duur kleiner wordt en dat de gemiddelde prestaties van de doelgroepleerlingen dichter in de buurt van het landelijk gemiddelde komen te liggen. In dit artikel wordt volstaan met de resultaten van de eerste meting. Hoewel het natuurlijk mogelijk is dat de effecten van het beleid pas op de lange termijn doordringen naar leerlingen, is het toch ook zo dat het OVB geen volledig nieuw beleid is, maar een voortzetting van het vroegere stimuleringsbeleid. Het doel van dat beleid was hetzelfde als het doel van het OVB, alleen de uitvoering van het beleid was anders.

Om enig zicht te krijgen op de werking van het beleid zullen in dit artikel tevens verschillen worden onderzocht tussen leerlingen die op verschillende manieren bij het beleid zijn betrokken. Gekeken wordt naar leerlingen van scholen die alleen extra basisformatie krijgen, leerlingen van scholen die in een onderwijs-

* Dit artikel is gebaseerd op het rapport van Tesser, P., Mulder, L. & Werf, M. P. C van der (1991). *De eerste fase van de longitudinale OVB-onderzoeken.* Nijmegen/Groningen: ITS/RION.



Figuur 1. Multi-level model voor het verklaren van prestatie-verschillen (vereenvoudigd op het schoolniveau)

voorrangsgebied participeren en leerlingen die op scholen zitten die geen of nauwelijks OVB-faciliteiten krijgen.

1 Onderzoeksmodel en vraagstellingen

Ten behoeve van de evaluatie van het OVB is een multi-level model ontwikkeld waarin twee niveaus zijn onderscheiden, namelijk het schoolniveau en het klasniveau. In het artikel van Van der Werf en Weide is het model vereenvoudigd op het leerlingniveau weergegeven. In Figuur 1 wordt het model vereenvoudigd op het schoolniveau gepresenteerd.

Op het leerlingniveau gaat het model ervan uit dat de invloed van de sociaal-etnische achtergrond van leerlingen op de prestaties in elk leerjaar indirect verloopt via het opvoedings- en culturele klimaat thuis dat weer van invloed is op het cognitieve beginniveau (intelligentie). Wat het OVB beoogt is de sterkte van deze invloed te doen afnemen, hetzij door maatregelen die gericht zijn op het thuisklimaat en/of op de cognitieve ontwikkeling.

Het bovenste deel van het model geeft de mogelijke werking van de verschillende vormen van betrokkenheid bij het OVB weer. Verondersteld wordt dat de relaties tussen de kenmerken van leerlingen en hun prestaties afhankelijk zijn van de betrokkenheid van scholen bij het OVB. Deze betrokkenheid kan ook invloed hebben op de prestaties van alle leerlingen op de scholen. Dan zijn de onderwijsachterstanden van de doelgroepleerlingen binnen deze scholen even groot als op andere scholen, maar absoluut gezien halen de doelgroepleerlingen daar dan hogere prestaties. In dat geval verminderen landelijk gezien de onderwijsachterstanden wel enigszins.

Uitgaande van het hierboven geschetste model zullen in dit artikel de volgende vraagstellingen worden beantwoord:

1 hoe sterk in het verband tussen sociaal-etnische achtergrond en leerprestaties?
2 hoe sterk is het verband tussen intelligentie en leerprestaties?

3 hoe sterk is het verband tussen thuisklimaat en leerprestaties?
4 hoe sterk is het verband tussen sociaal-etnische achtergrond en leerprestaties, voor en na controle voor intelligentie en thuisklimaat?
5 zijn er verschillen tussen schooltypen (groepen scholen die op verschillende wijze participeren in het OVB) in de leerprestaties van de doelgroepen van het OVB voor en na controle voor intelligentie en thuisklimaat?

2 Methode

2.1 Steekproef en respons

De algemene opzet van de evaluatie en de steekproefopzet worden vrij uitvoerig besproken in het artikel van Van der Werf en Weide en in het artikel van Tesser en Vierke in dit nummer. Samengevat komt de opzet erop neer dat om de twee jaar toetsen worden afgenomen bij alle leerlingen van groep 4, 6 en 8 van de basisscholen die in de steekproef zijn opgenomen. De steekproef is gestratificeerd naar de gebiedenfactor (wel - geen gebiedsschool) en naar de schoolscorefactor (<105; 106-115; >115). De cel gebiedsscholen met een schoolscore <105 is leeg, zodat er uiteindelijk vijf groepen scholen resteren. Voor de overzichtelijkheid wordt in dit artikel een indeling in drie groepen gehanteerd: scholen zonder OVB-faciliteiten, scholen met alleen extra basisformatie en gebiedsscholen. Daarnaast is er een referentiegroep, de 3,5% aselecte steekproef uit alle basisscholen. De oorspronkelijke steekproef bestond uit in totaal 1237 scholen, die alle benaderd zijn met het verzoek om medewerking te verlenen aan het onderzoek. De respons bedroeg 56%. In Tabel 1 staat de respons per groep scholen weergegeven.

Tabel 1
Aantallen benaderde en deelnemende scholen per groep

groep	benaderd N	respons N	(%)
Scholen zonder OVB-faciliteiten	329	140	(43)
Scholen met alleen extra basisformatie	426	183	(43)
Gebiedsscholen	482	373	(77)

De respons van de gebiedsscholen is aanmerkelijk hoger dan die van de beide andere groepen scholen. De gebiedsscholen zijn niet helemaal representatief voor de totale populatie gebiedsscholen. Gebiedsscholen in de grote steden zijn duidelijk ondervertegenwoordigd. Daarmee samenhangend zijn er ook verschillen in richting (meer katholieke scholen in de steekproef) en naar schoolscore (minder scholen met een schoolscore >115). Algemene uitspraken over gebiedsscholen dienen dus met de nodige voorzichtigheid te geschieden. Wat betreft de referentiegroep, deze is wel representatief voor alle basisscholen, zowel qua kenmerken van de deelnemende scholen als qua weegfactor van de leerlingen in de referentiegroep.

2.2 Variabelen en instrumenten

In de vraagstellingen van dit artikel zijn de volgende variabelen aan de orde: sociaal-etnische achtergrond van de leerlingen, cognitieve ontwikkeling, thuisklimaat en leerprestaties. De variabelen cognitieve ontwikkeling en leerprestaties zijn gemeten met respectievelijk intelligentietesten ontwikkeld door het RION en met taal- en rekentoetsen ontwikkeld door het CITO. De intelligentietesten bestaan voor alle groepen uit vijf subtesten, twee voor 'ruimtelijk inzicht', twee voor 'semantische vaardigheden', en één voor 'symbolisch inzicht'. Voor dit artikel worden alleen de twee subtesten voor ruimtelijk inzicht gebruikt, omdat dit aspect van de intelligentie verondersteld wordt het meest het begrip 'aanleg' te meten.

De taaltoetsen meten het niveau van de algemene taalvaardigheid die alle leerlingen in de respectievelijke groepen op het moment van de afname zouden moeten beheersen. De toets bestaat uit morfologische, syntactische en semantische opgaven.

De rekentoetsen meten eveneens 'minimale communale doelen', en bestaan uit opgaven met betrekking tot getallen, ordenen, automatiseren, bewerkingen en tellen.

De sociaal-etnische achtergrond van de leerlingen is geïndiceerd door de weegfactor die de school van elke leerling heeft verstrekt. Een weegfactor van 1.00 geeft aan dat de leerling een Nederlandse achtergrond heeft en afkomstig is uit een midden tot hoog milieu. Een 1.25 leerling is een Nederlandse leerling af-

komstig uit een laag milieu en een 1.90 leerling is een allochtone leerling uit een laag milieu. Van de laatstgenoemde groep leerlingen is er daarnaast informatie over het geboorteland van de ouders. Behalve gegevens over de weegfactor en de etnische herkomst is er ook informatie verzameld over de opleiding en het beroepsniveau van beide ouders van de leerlingen. Hiertoe zijn vragenlijsten aan de leerlingen verstrekt met het verzoek deze door de ouders te laten in vullen. Omdat de non-respons op deze vragenlijsten zeer hoog was en tevens selectief, moeten deze gegevens in dit artikel met de nodige voorzichtigheid worden geïnterpreteerd. Dit geldt eveneens voor de gegevens over het 'onderwijsondersteunend gedrag' en 'culturele hulpbronnen' van de ouders die ook verkregen zijn met de oudervragenlijst. Voor de wijze waarop deze begrippen zijn geoperationaliseerd wordt verwezen naar de paragraaf die handelt over resultaten met betrekking tot het thuisklimaat en de leerprestaties van de leerlingen.

3 Resultaten

3.1 De relatie tussen sociaal-etnische achtergrond en leerprestaties

In Tabel 2 zijn de gemiddelde scores van de leerlingen op de toetsen per gewichtscategorie weergegeven. De scores zijn standaard normaalscores (T-scores) met een gemiddelde van 50 en een standaarddeviatie van 10, berekend op basis van de verdeling van de toetsresultaten van de leerlingen in de referentiesteekproef. Voor de overzichtelijkheid zijn in deze tabel de standaarddeviaties voor de onderscheiden groepen weggelaten. Ze variëren tussen 7 en 10.

Globaal is het beeld in elke jaargroep ongeveer hetzelfde. De 1.00 leerlingen scoren zowel voor taal als voor rekenen iets boven het gemiddelde van de referentiegroep. De 1.25 leerlingen scoren daar iets onder. De 1.90 leerlingen scoren ver onder het gemiddelde van de referentiegroep. Vergelijken we de scores van de leerlingen over de jaargroepen, dan zien we dat de verschillen tussen de onderwijsachterstanden van de 1.90 leerlingen groter lijken te worden naarmate ze langer onderwijs hebben gevolgd. Vooral bij de taalscores is dit het geval. In groep 4 scoren de 1.90 leerlingen iets meer dan een halve standaarddeviatie onder het referentiegemiddelde, in groep 6 bedraagt het verschil al bijna een hele standaarddeviatie. In vergelijking met de 1.00 leerlingen wordt het beeld nog pregnanter: terwijl de scores van de 1.90 leerlingen met het toenemen van het leerjaar steeds lager worden, blijven die van de 1.00 leerlingen gelijk. Voor rekenen worden de scores van de laatstgenoemde groep zelfs hoger. Hetzelfde zien we in enige mate in de vergelijking tussen de 1.25 en de 1.00 leerlingen. In groep 4 zijn deze verschillen nog klein, maar in groep 6 en 8 zijn deze toegenomen tot bijna een halve standaarddeviatie. Op grond van deze gegevens kan geconcludeerd worden dat het verband tussen sociaal-etnische achtergrond en schoolprestaties behoorlijk sterk is.

Tabel 2

Gemiddelde T-scores voor taal en rekenen, per wegingsfactor en per jaargroep (1.00 en 1.25 leerlingen: referentiegroep; 1.90 leerlingen: totale steekproef)

		Taal		Rekenen	
		Gemiddelde	*Aantal ll.*	*Gemiddelde*	*Aantal ll.*
Groep 4	1.00	53	2520	51	2441
	1.25	51	1846	49	1789
	1.90	44	3071	43	3052
Groep 6	1.00	52	2508	52	2502
	1.25	48	1864	48	1860
	1.90	41	2825	43	2814
Groep 8	1.00	53	2509	52	2521
	1.25	49	1945	48	1927
	1.90	39	2544	42	2542

3.2 De relatie tussen cognitieve ontwikkeling en de leerprestaties

In feite kunnen we pas spreken van onderwijsachterstanden als de prestaties van bepaalde groepen leerlingen, die qua aanleg vergelijkbaar zijn met andere groepen, desondanks lager zijn dan die van die andere groepen. Dit is de reden dat bij het analyseren van verschillen in prestaties tussen groepen gecontroleerd moet worden voor aanleg. Voor het meten van aanleg moeten we noodgedwongen volstaan met een intelligentietest, die helaas echter ook altijd aangeleerde schoolse danwel niet-schoolse vaardigheden meet. Om het begrip aanleg zo zuiver mogelijk onder controle te houden gebruiken we alleen die onderdelen van de intelligentietest die zo weinig mogelijk een beroep doen op verbale, aangeleerde vaardigheden. Dit zijn de onderdelen 'exclusie' en 'figuur samenstellen' van de intelligentietest die speciaal voor de evaluatie OVB is ontwikkeld.

In Tabel 3 wordt eerst een overzicht gegeven van de scores van de leerlingen op deze testonderdelen. Evenals bij de taal- en rekenscores zijn de intelligentiescores uitgedrukt in T-scores met een gemiddelde van 50 en een standaarddeviatie van 10. De standaarddeviaties van de onderscheiden groepen variëren tussen 8 en 10.

Tabel 3

Gemiddelde T-scores op het non-verbale deel van de intelligentietest, per jaargroep en per wegingsfactor (1.00 en 1.25 leerlingen: referentiegroep; 1.90 leerlingen: totale steekproef)

	Wegingsfactor		
	1.00	1.25	1.90
Groep 4	52	48	43
Groep 6	51	49	45
Groep 8	52	49	47

Wat opvalt in deze tabel is dat de achterstanden van de 1.90 leerlingen ten opzichte van de 1.00 en 1.25 leerlingen minder groot zijn dan die voor taal en rekenen. Verder valt op dat de achterstanden gedurende de basisschoolperiode kleiner worden (van bijna een hele standaardafwijking in groep 4 naar een halve standaardafwijking in groep 8), dit eveneens in tegenstelling tot de achterstanden in taal- en rekenprestaties die in de loop van de tijd juist toenemen.

In Tabel 4 zijn de correlaties tussen intelligentie en de taal- en rekenprestaties weergegeven. De samenhangen zijn redelijk sterk, zodat mag worden aangenomen dat bij controle voor intelligentie de achterstanden van de doelgroepleerlingen van het OVB heel wat kleiner zullen worden.

Tabel 4

Pearson correlaties tussen intelligentie en taal- en rekenscores, per jaargroep (totale steekproef)

	Taal	Rekenen
Groep 4	.49	.57
Groep 6	.40	.50
Groep 8	.42	.52

De correlaties met de rekenscores zijn hoger dan die met de taalscores. Verder valt op dat de verbanden in leerjaar 4 sterker zijn dan in de leerjaren 6 en 8. Op grond van deze gegevens mag worden aangenomen dat bij controle voor intelligentie het effect van de wegingsfactor kleiner zal worden en het sterkst zal afnemen in leerjaar 4. Of deze aanname correct is zullen we verderop zien.

3.3 Relaties tussen thuisklimaat en leerprestaties

De variabelen die betrekking hebben op het thuisklimaat van leerlingen zijn ingedeeld in twee categorieën: de beschikbaarheid van culturele hulpbronnen in het gezin en schoolondersteunend gedrag van ouders. Naar voorbeeld van De Graaf (1987) zijn met betrekking tot de beschikbaarheid van culturele hulpbronnen twee typen onderscheiden, namelijk 1) boeken hebben, lezen, kopen en 2) cultuurparticipatie in de zin van concert-, theater- en museumbezoek. De gegevens over 'boeken' en 'cultuurparticipatie' zijn bij de ouders verzameld met behulp van schriftelijke vragenlijsten. Na factoranalyse is voor elke variabele een schaal geconstrueerd, elk bestaande uit zes items, met een betrouwbaarheid van respectievelijk .80 en .87.

Ook de gegevens over schoolondersteunend gedrag van ouders zijn verzameld met schriftelijke vragenlijsten. Na factoranalyse op de 12 items resulteren twee schalen: 1) 'positieve communicatie' bestaande uit zes items met een betrouwbaarheid van .81 en 2) 'reactie op slechte prestaties', bestaande uit vier items met een betrouwbaarheid van .75.

In Tabel 5 staan de gemiddelde scores op de vier schalen per wegingsfactor.

Tabel 5

Gemiddelde scores op de thuisklimaat variabelen, per wegingsfactor over de drie cohorten tezamen (max. score is 4). 1.00 en 1.25 leerlingen: referentiesteekproef; 1.90 leerlingen: totale steekproef

	Wegingsfactor		
	1.00	1.25	1.90
Boeken	2.6	2.2	2.0
Cultuurparticipatie	1.7	1.3	1.2
Positieve communicatie	3.7	3.7	3.4
Reactie op prestaties	2.3	2.4	2.8

Uit deze tabel blijkt dat de ouders van de 1.25 en 1.90 leerlingen in mindere mate beschikken over culturele hulpbronnen dan de ouders van de 1.00 leerlingen. Wat betreft cultuurparticipatie valt tevens op dat ook de ouders van 1.00 leerlingen tamelijk laag scoren. Ouders van 1.9 leerlingen praten minder vaak met hun kinderen over school dan ouders van 1.25 en 1.00 leerlingen. Tussen de laatstgenoemde twee groepen is geen verschil. Ten slotte blijkt dat de ouders van de 1.00 leerlingen minder reageren op slechte prestaties dan die van 1.25 leerlingen en deze ouders doen dat weer minder dan ouders van 1.90 leerlingen. Of dit laatste te maken heeft met het feit dat 1.25 en 1.90 leerlingen ook slechter presteren valt af te leiden uit Tabel 6. In deze tabel zijn de correlaties weergegeven tussen elk van de thuisklimaatvariabelen en de prestaties voor taal en rekenen.

Het patroon van de correlaties is in alle leerjaren vrijwel gelijk. De correlaties tussen boeken en culturele participatie enerzijds en prestaties anderzijds zijn voor taal hoger dan voor rekenen. Voor zowel taal als rekenen geldt ook dat het verband tussen prestaties en de beide culturele hulpbronnenvariabelen met het leerjaar toeneemt. Dit wijst enigszins op een mogelijke causale invloed van de beschikbaarheid van culturele hulpbronnen op de prestaties.

De correlaties tussen positieve communicatie en de prestaties zijn zeer zwak, maar ook voor taal iets hoger dan voor rekenen. Ook hier nemen de verbanden in sterkte toe met het leerjaar. Ten slotte zijn de correlaties tussen reactie op slechte prestaties en de prestaties zelf negatief. Dit wijst er op dat reacties op slechte prestaties meer gevolg dan oorzaak zijn van prestaties.

De consequentie van deze bevindingen is dat voor het analyseren van de relatieve bijdrage van thuisklimaatvariabelen op de onderwijsachterstanden van de doelgroepen van het OVB voor dit onderzoek alleen de beschikbaarheid van culturele hulpbronnen relevant is.

3.4 Relatieve invloed van sociaal-etnische achtergrond, intelligentie en thuisklimaat

In het voorgaande is geconstateerd dat er een sterk effect is van de sociaal-etnische achtergrond van leerlingen op hun prestaties, alsmede dat er een zeer sterk verband is tussen intelligentie en prestaties en een zwak verband tussen beschikbaarheid van culturele hulpbronnen en prestaties. In het vervolg zal worden nagegaan wat er van het effect van de sociaal-etnische achtergrond van leerlingen overblijft na controle voor intelligentie en culturele hulpbronnen. Alleen als na deze controle de prestaties van de doelgroepen van het OVB nog substantieel van elkaar verschillen, mag worden gesproken van een onderwijsachterstand van deze groepen.

In Tabel 7 staan de effecten van de wegingsfactoren zonder en met controle voor intelligentie.

In alle gevallen worden de effecten van de wegingsfactor kleiner na controle voor intelligentie. De effecten na controle zijn in groep 4 wat kleiner dan in groep 6 en 8. Dit bevestigt het eerdere vermoeden dat de onderwijsachterstanden in de loop van de basisschoolperiode eerder toe dan afnemen, terwijl de intelli-

Tabel 6

Correlaties tussen thuisklimaatvariabelen en de toetsscores voor taal en rekenen, per leerjaar (totale steekproef)

	Groep 4		Groep 6		Groep 8	
	Taal	Rekenen	Taal	Rekenen	Taal	Rekenen
Boeken	.21	.15	.23	.16	.24	.18
Participatie	.19	.14	.22	.17	.24	.20
Positieve communicatie	.09	.05	.11	.06	.13	.08
Reactie op prestaties	-.23	-.21	-.27	-.26	-.24	-.24

Tabel 7

Ongestandaardiseerde regressie-effecten van de taal- en rekenscores in groep 4, 6 en 8 op de wegingsfactoren 1.25 en 1.90 voor en na controle voor intelligentie (totale steekproef)

	Taal			Rekenen		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
Wegingsfactor 1.25						
voor controle	-4.5	-5.1	-5.3	-3.5	-4.8	-5.1
na controle	-2.7	-4.0	-4.0	-1.1	-3.3	-3.3
Wegingsfactor 1.90						
voor controle	-11.3	-11.3	-12.6	-7.6	-9.1	-9.6
na controle	-8.0	-9.3	-10.5	-3.3	-6.4	-6.9

Alle effecten zijn significant ($p < .01$)

Tabel 8

Ongestandaardiseerde regressie-effecten van de taal- en rekenscores in groep 4, 6 en 8 op de wegingsfactoren 1.25 en 1.90 voor en na controle voor culturele hulpbronnen (totale steekproef)

	Taal			Rekenen		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
Wegingsfactor 1.25						
voor controle	-4.5	-5.1	-5.3	-3.5	-4.8	-5.1
na controle	-3.5	-3.8	-3.8	-2.7	-4.1	-4.1
Wegingsfactor 1.90						
voor controle	-11.3	-11.3	-12.6	-7.6	-9.1	-9.6
na controle	-10.2	-9.9	-11.0	-6.9	-8.4	-8.6

Alle effecten zijn significant ($p < .01$)

gentie toeneemt. Verder is in deze tabel te zien dat het effect van de controle sterker is voor de 1.90 leerlingen dan voor de 1.25 leerlingen, hetgeen er wellicht op wijst dat de prestaties van de 1.25 leerlingen meer in overeenstemming zijn met hun aanleg dan die van de 1.90 leerlingen.

In Tabel 8 staan de effecten van de wegingsfactoren zonder en met controle voor culturele hulpbronnen.

Ook hier worden de effecten van de wegingsfactoren in alle gevallen kleiner na controle voor culturele hulpbronnen. De afname van de effecten is echter beduidend kleiner dan bij controle voor intelligentie. Bij de interpretatie van deze effecten dient echter rekening te worden gehouden met de hoge mate van (selectieve) non-respons op de vragenlijsten waarmee de betreffende variabele gemeten is. Tevens kunnen vraagtekens worden gezet bij de validiteit van de vragenlijst.

3.5 Verschillen tussen schooltypen

De schooltypen die in dit artikel worden onderscheiden, namelijk gebiedsscholen, scholen met alleen extra basisformatie (formatiescholen) en scholen zonder OVB-faciliteiten representeren in het onderzoek de middelen die het

Tabel 9

Gemiddelde T-scores op taal en rekenen per schooltype in groep 4, 6 en 8 (totale steekproef)

		Gebiedsscholen	Formatiescholen	Scholen zonder OVB-faciliteiten
Taal	groep 4	44	49	51
	groep 6	45	49	51
	groep 8	44	48	52
Rekenen	groep 4	46	49	51
	groep 6	46	48	52
	groep 8	45	48	52

OVB hanteert om de doelstellingen van het beleid te bereiken. De prestaties van de leerlingen en speciaal de prestaties van de leerlingen uit de doelgroepen van het beleid op de gebiedsscholen en de formatiescholen zeggen iets over de effecten van het beleid.

In Tabel 9 worden de gemiddelde T-scores op de taal- en rekentoetsen van de leerlingen op de verschillende schooltypen gepresenteerd.

De gebiedsscholen scoren 4 tot 6 punten onder het referentiegemiddelde, de formatiescholen 1,5 tot 2 punten. De scores van de scholen zonder OVB-formatie liggen 1 tot 1,5 punt boven het referentiegemiddelde. Voor een juiste interpretatie zal echter rekening moeten worden gehouden met de samenstelling van de leerlingpopulatie van de schooltypen omdat vooral op de gebiedsscholen zeer veel allochtone leerlingen zitten die ver onder het referentiegemiddelde scoren (Tabel 2). In Tabel 10 worden daarom de scores van de leerlingen op de verschillende schooltypen per wegingsfactor vergeleken.

We zien dat de toetsscores van alle drie de groepen leerlingen op de gebiedsscholen lager liggen dan op de formatiescholen en dat ze daar weer lager liggen dan op de scholen zonder OVB-faciliteiten. Een verklaring hiervoor kan zijn dat de sociaal-etnische achtergrond van de leerlingen op deze scholen ongunstiger is. Dit ondanks het feit dat reeds voor de wegingsfactor is gecontroleerd. Deze wegingsfactor is echter een zeer grove indicator van sociaal-etnische achtergrond, zodat binnen de wegingsfactoren nog aanmerkelijke verschillen kunnen bestaan. Omdat van een groot deel van de leerlingen ook meer precieze informatie beschikbaar is over de opleiding en het beroepsniveau van de ouders, alsmede over het herkomstland van allochtone leerlingen is opnieuw een analyse uitgevoerd van de verschillen tussen schooltypen in toetsprestaties van de groepen leerlingen met controle voor beroep en opleiding van de vader, opleiding van de moeder en herkomstland. In Tabel 11 staan de resultaten van de regressie-analyse waarin voor elke wegingsfactor apart de effecten van schooltypes zijn geschat na controle voor de sociale-etnische kenmerken. Schooltype is in deze analyses opgenomen als twee dummy variabelen, één voor gebiedsscholen en één voor formatiescholen.

Tabel 11

Ongestandaardiseerde regressie effecten van de toetsscores voor taal en rekenen op schooltype (G=gebiedsscholen, F=formatiescholen) na controle voor sociaal-etnische achtergrond kenmerken, per jaargroep en per wegingsfactor

		Taal		Rekenen	
		G	F	G	F
Groep 4	1.00	-1.1*	-0.1*	-1.1*	-0.7*
	1.25	-3.3	-1.7	-2.4	-1.4*
	1.90	-0.6*	-1.8	-1.8*	-1.6*
Groep 6	1.00	-0.9*	-0.3*	-1.5	-1.2*
	1.25	-2.1	-0.4*	-2.6	-1.6*
	1.90	-2.3*	-1.2*	-2.2*	-2.2*
Groep 8	1.00	-1.6	-0.5*	-1.5	-0.9*
	1.25	-3.8	-1.9	-3.2	-2.4
	1.90	-0.6*	-0.2*	-3.6*	-3.3*

* = niet significant ($p > .01$)

Uit deze tabel blijkt dat de resterende verschillen voor de 1.00 leerlingen erg klein zijn, maar soms nog wel statistisch significant (voor taal

Tabel 10

Gemiddelde T-scores op taal en rekenen per jaargroep en wegingsfactor op de verschillende schooltypen (totale steekproef); G=gebiedsscholen, F=formatiescholen en ZF=scholen zonder OVB-faciliteiten

		Taal			Rekenen		
		G	F	ZF	G	F	ZF
Groep 4	1.00	50	52	52	50	51	52
	1.25	46	48	50	47	48	50
	1.90	40	43	42	43	44	47
Groep 6	1.00	50	52	53	51	51	53
	1.25	46	48	49	46	47	49
	1.90	40	42	45	42	44	47
Groep 8	1.00	51	52	53	51	51	53
	1.25	45	48	50	46	47	50
	1.90	39	41	43	42	43	48

in groep 8: voor rekenen in groep 6 en 8). Het gaat hierbij alleen om effecten van de gebiedsscholen.

Voor de 1.25 leerlingen geldt voor de effecten van de gebiedsscholen dat ze nog steeds significant zijn, zowel voor taal als voor rekenen in alle jaargroepen. Het gaat tevens nog steeds om behoorlijke verschillen. De effecten van de formatiescholen zijn voor de 1.25 leerlingen kleiner dan die voor de gebiedsscholen en nog slechts in drie gevallen significant (voor taal in groep 4 en 8 en voor rekenen in groep 8).

De effecten van de gebiedsscholen voor de 1.90 leerlingen zijn in geen enkel geval statistisch significant (vanwege de kleinere aantallen), maar qua grootte zijn de effecten in een aantal gevallen vergelijkbaar met die voor de 1.25 leerlingen (voor taal in groep 6 en voor rekenen in groep 6 en 8). In de andere gevallen zijn de effecten klein tot zeer klein. De effecten van de formatiescholen zijn slechts in één geval significant (voor taal in groep 4). In sommige gevallen zijn ze vergelijkbaar met die voor de 1.25 leerlingen (voor taal in groep 4 en voor rekenen in groep 4), soms zijn ze groter (voor taal in groep 6, voor rekenen in groep 6 en 8) en soms zijn ze kleiner (voor taal in groep 8). Echt duidelijk is het beeld dus niet. Wanneer de effecten van de gebiedsscholen en formatiescholen worden vergeleken wordt het beeld niet duidelijker. De meeste verschillen zijn klein. De verschillen die er wel zijn, zijn niet consistent in dezelfde richting: voor taal in groep 4 is het effect sterker voor de formatiescholen; voor taal in groep 6 zijn ze sterker voor de gebiedsscholen.

Voorlopig kan uit de analyses alleen geconcludeerd worden dat de prestaties van de 1.25-leerlingen op de gebiedsscholen, ook na controle voor hun sociale achtergrond, duidelijk achterblijven bij die van de 1.25-leerlingen op de scholen zonder OVB-faciliteiten, maar ook bij die van de 1.25-leerlingen op de formatiescholen. Er dient echter nogmaals te worden gewezen op de grote (selectieve) uitval op de oudervragenlijsten waarmee de sociale achtergrondkenmerken zijn verzameld.

Theoretisch is het ook mogelijk dat het intelligentie niveau van de 1.25-leerlingen op de gebiedsscholen lager ligt dan van de 1.25-leerlingen op de formatiescholen en de scholen zonder OVB-faciliteiten. Ook dit zou de lagere prestaties van de 1.25 leerlingen op de gebiedsscholen kunnen verklaren. Daarom is nogmaals een regressie-analyse uitgevoerd voor het vaststellen van de effecten van de schooltypen op de prestaties van de 1.25-leerlingen, na controle voor intelligentie en sociale achtergrond gezamenlijk. De resultaten staan in Tabel 12.

Tabel 12

Ongestandaardiseerde regressie-effecten van taal- en rekenscores van 1.25 leerlingen op formatie- en gebiedsscholen na controle voor sociale achtergrond en intelligentie (totale steekproef)

	Taal		Rekenen	
	G	F	G	F
Groep 4	-2.5	-1.1	-1.3*	-0.7*
Groep 6	-1.4*	0.0*	-1.4	-1.0*
Groep 8	-3.3	-1.6*	-2.1	-2.5

* = niet significant ($p > .01$)

Bij de gebiedsscholen zijn nog 3 van de 6 effecten voor 1.25 leerlingen significant, bij de formatiescholen eveneens nog 3.

De effecten van de formatiescholen zijn in alle gevallen, behalve bij rekenen in groep 8, kleiner dan de effecten van de gebiedsscholen. De effecten van de gebiedsscholen zijn voor taal sterker dan voor rekenen. Op de verklaring van deze bevindingen wordt in de conclusies nader ingegaan.

4 Conclusies

Bij het trekken van conclusies uit de hierboven gepresenteerde gegevens moet opgemerkt worden dat over de effecten van het OVB-beleid, zoals dat vanaf 1986 is gevoerd, vooralsnog niet veel valt te zeggen. De gegevens hebben immers slechts betrekking op één meetmoment. Dat neemt echter niet weg dat wel iets gezegd kan worden over de situatie anno 1988. We zullen dit doen aan de hand van de vijf vraagstellingen die in dit artikel zijn onderzocht.

De eerste vraagstelling betreft de vraag naar de relatie tussen de sociaal-etnische kenmerken van leerlingen en hun prestaties. De resultaten leveren een eenduidig beeld op. Zowel de taal- als de rekenprestaties van de allochtone leerlingen blijven in alle drie onderzochte jaargroepen ver achter bij het landelijk gemid-

delde. Binnen de groep allochtone leerlingen zijn er weliswaar duidelijke verschillen in samenhang met herkomstland, maar ook de best presterende 1.90-leerlingen halen bij lange na niet het landelijk gemiddelde. Bij de Nederlandse 1.25-leerlingen zijn de resultaten wat minder eenduidig. Hun gemiddelde scores liggen onder het landelijk gemiddelde, maar de verschillen zijn niet zo groot.

De tweede vraagstelling is gericht op de relatie tussen intelligentie en taal- en rekenprestaties. Deze relaties blijken redelijk sterk te zijn, vooral de relaties tussen intelligentie en rekenprestaties. Opvallend is dat de verbanden afnemen naarmate leerlingen meer jaren basisonderwijs hebben gevolgd. Wat ook opvalt is dat de achterstand in intelligentie van de doelgroepleerlingen met het leerjaar afneemt, dit in tegenstelling tot de achterstand in prestaties, die met het leerjaar toeneemt. Dit zou erop kunnen wijzen dat gedurende de schoolperiode achterstandsleerlingen steeds meer onder hun niveau (qua aanleg) gaan presteren.

De derde vraagstelling gaat over de relaties tussen het thuisklimaat en taal- en rekenprestaties. Thuisklimaat is onderscheiden in beschikbaarheid van culturele hulpbronnen en onderwijsondersteunend gedrag van ouders, die op hun beurt weer zijn onderverdeeld in 'boeken' en 'cultuurparticipatie', respectievelijk 'positieve communicatie' en 'reactie op slechte prestaties'. Ouders van allochtone leerlingen beschikken minder over culturele hulpbronnen en praten minder vaak met hun kinderen over school. Tussen ouders van 1.00- en 1.25-leerlingen zijn de verschillen klein. Wat betreft reacties op slechte prestaties blijkt dat ouders van 1.90-leerlingen hierop hoger scoren. De correlaties tussen de vier thuisklimaatvariabelen en taal- en rekenprestaties geven een wisselend beeld. De correlaties tussen 'boeken' en 'cultuurparticipatie' zijn zwak (.14-.24), voor taal hoger dan voor rekenen en ze nemen toe met het leerjaar.

De correlaties tussen positieve communicatie en prestaties zijn zeer zwak, maar ook voor taal iets hoger dan voor rekenen. De correlaties tussen reactie op slechte prestaties en de prestaties zelf zijn zwak, maar tevens negatief. Dit wijst erop dat reacties op slechte prestaties meer gevolg dan oorzaak zijn van prestaties.

Vraagstelling 4 betreft de vraag naar de relatieve invloed van sociaal-etnische achtergrond, intelligentie en thuisklimaat. Nagegaan is wat er van het effect van sociaal-etnische achtergrond op de prestaties overblijft na controle voor intelligentie en culturele hulpbronnen. Uit de resultaten blijkt dat de effecten van de wegingsfactor na controle voor intelligentie in alle gevallen kleiner wordt. Dit is het meest het geval in groep 4. Dit bevestigt het eerdere vermoeden, dat de onderwijsachterstanden van de doelgroepleerlingen in de loop van de basisschoolperiode eerder toe- dan afnemen, ook na controle voor intelligentie. Na controle voor de beschikbaarheid van culturele hulpbronnen wordt het effect van de wegingsfactoren eveneens kleiner, maar aanmerkelijk minder dan bij controle voor intelligentie.

De laatste vraagstelling is gericht op verschillen tussen typen scholen die op verschillende manier deelnemen aan het OVB. Onderscheiden zijn gebiedsscholen, formatiescholen en scholen zonder OVB-faciliteiten.
Uit de eerste analyses blijkt dat de gebiedsscholen fors lager scoren dan de formatiescholen en deze weer lager dan de scholen zonder OVB-faciliteiten. Het is echter mogelijk dat de samenstelling van de schoolbevolking aan deze verschillen debet is. Op de gebiedsscholen zitten namelijk veel meer allochtone leerlingen dan op de beide groepen andere scholen. Vergelijking tussen de drie typen scholen per wegingsfactor laat echter zien dat ook bij gelijke wegingsfactor de gebiedsscholen nog steeds lager scoren dan de formatiescholen en deze weer lager dan de scholen zonder OVB-formatie. Daarom is, ondanks dat van slechts een deel van de leerlingen deze informatie beschikbaar is, ook nog eens gecontroleerd voor het opleidingsniveau van de ouders en het herkomstland van leerlingen.

Na deze controle blijven er alleen voor de 1.25-leerlingen nog steeds behoorlijke effecten van de gebiedsscholen over. Als ook nog gecontroleerd wordt voor intelligentie, blijven de meeste van deze effecten nog steeds significant. De conclusie die hieruit getrokken zou kunnen worden, is dat er een negatief effect is van deelname aan onderwijsvoorrangsgebieden op de prestaties van Nederlandse arbeidersleerlingen. Deze conclusie is echter vooralsnog voorbarig. Enerzijds omdat nog slechts één meetmoment is gerealiseerd, ander-

zijds omdat er aanwijzingen zijn dat de negatieve effecten van de gebiedsscholen toegeschreven kunnen worden aan het hoge percentage 1.90-leerlingen op die scholen. Uit de analyses van Tesser, Mulder en Van der Werf (1991) en Tesser en Mulder (1990) blijkt dat de achterstanden van de gebiedsscholen nagenoeg verdwijnen als rekening wordt gehouden met het percentage 1.90 leerlingen op een school. Merkwaardig genoeg is dit dus na controle van de individuele achtergrondkenmerken van de leerlingen. Er lijkt dus sprake van een context-effect, namelijk van een effect van de samenstelling van de schoolbevolking op de prestaties van leerlingen dat bovenop het effect komt van de individuele sociaal-economische kenmerken van de leerlingen. In hoeverre het context-effect eigenlijk een schooleffect is, in de zin dat het onderwijs op scholen met een hoog percentage 1.90-leerlingen minder effectief georganiseerd is, zal verder worden uitgezocht op basis van de gegevens van de tweede en derde meting van de evaluatie van het OVB.

Literatuur

Tesser, P., Mulder, L., & Werf, M. P. C. van der (1991). *De eerste fase van de longitudinale OVB-onderzoeken: het leerlingenonderzoek*. Nijmegen/ Groningen: ITS/RION.

Graaf, P. M. de (1987). *De invloed van financiële en culturele hulpbronnen in onderwijsloopbanen*. Nijmegen: ITS.

Tesser, P., & Mulder, L. (1990). Concentratie en prestaties van allochtone leerlingen in het basisonderwijs. *Migrantenstudies, 6*,(2), 31-44.

Manuscript aanvaard 11-1-1993

Auteurs

M. P. C. van der Werf is werkzaam bij het RION, Instituut voor Onderwijsonderzoek van de Rijksuniversiteit Groningen als projectleider van de landelijke evaluatie van het OVB.

P. Tesser werkte tot 1992 bij het Instituut voor Toegepaste Sociale Wetenschappen (ITS) te Nijmegen en was daar onder andere projectleider van de Landelijke Evaluatie Onderwijsvoorrangsbeleid. Is sinds juni 1992 werkzaam bij het Sociaal Cultureel Planbureau.

L. Mulder is werkzaam bij het Instituut voor Toegepaste Sociale Wetenschappen (ITS), te Nijmegen en is sinds 1987 betrokken bij de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid.

Correspondentie-adres: RION, Instituut voor Onderwijsonderzoek. Postbus 1286, 9701 BG Groningen.

Abstract

The achievements of the Educational Priority target groups in 1988; state of the art after the first measurement

M. P. C. van der Werf, P. Tesser & L. Mulder. Pedagogische Studiën, 1993, 70, 97-107.

The data of the first measurement of the evaluation of the Educational Priority Policy (OVB) in 1988 are analyzed to figure out the relationships between the social ethnic background, intelligence and family characterics of pupils and their achievements in arithmetic and Dutch language in grade 4, 6 and 8 of primary education. Also differences between categories of schools are analyzed. The language as well as the arithmetic achievements of immigrant pupils lay behind the mean scores of the national reference group. The disadvantages of the 1.25 pupils are much smaller. There appears a strong relationship with intelligence and a weak relationship with cultural resources at home. After correction for intelligence the effect of weight factors becomes importantly weaker, after correction for cultural resources this effect becomes only a bit weaker.

Schools participating in a priority area score much lower than schools having only general OVB facilities; these last ones score lower than schools without any OVB facilities. After correction for weight factor of the students, the educational level of their parents and their ethnic origin the effect of the area schools only remains for 1.25 pupils. This effect also remains after correction for intelligence. From other analyses it appears that this effect almost completely is explained by the high percentage of 1.90 student in area schools.

Effectieve voorrangskenmerken in de school en in de klas

M. P. C. van der Werf en M. G. Weide

Samenvatting

In de eerste cohortstudie van de landelijke evaluatie van het Onderwijsvoorrangsbeleid, uitgevoerd in 1988, is door middel van schriftelijke vragenlijsten, afgenomen bij directeuren en leerkrachten van basisscholen, informatie verzameld over school en klaskenmerken. Bij de leerlingen in groep 4, 6 en 8 zijn testen afgenomen voor intelligentie, taal en rekenen. Uit de resultaten blijkt dat er enige verschillen zijn in school- en klaskenmerken tussen scholen met verschillende aantallen achterstandsleerlingen, maar de verschillen zijn niet erg groot. Scholen met veel allochtone leerlingen scoren wat hoger op enkele school- en instructie-effectiviteitskenmerken en op enkele variabelen met betrekking tot specifieke strategieën voor achterstandsleerlingen vergeleken bij de andere groepen scholen.

Uit multi-level analyses blijkt echter dat bijna geen van deze kenmerken samenhangt met verschillen tussen scholen in de prestaties van leerlingen. Alleen de frequentie waarmee huiswerk wordt gegeven hangt tamelijk consistent samen met de rekenprestaties.

Inleiding

"Het Onderwijsvoorrangsbeleid is erop gericht door middel van een samenspel van gerichte maatregelen individuele scholen en welzijnsinstellingen, dan wel scholen en welzijnsinstellingen in een samenwerkingsverband verenigd, in staat te stellen onderwijsachterstanden van leerlingen van 4 tot 18 jaar ten gevolge van sociale, economische en culturele omstandigheden op te heffen of te verminderen" (Onderwijsvoorrangsplan p. 5), (Ministerie van Onderwijs en Wetenschappen, 1985). Uit deze formulering van de doelstelling van het Onderwijsvoorrangsbeleid (OVB) kan in de eerste plaats worden afgeleid dat het beleid een faciliteitenbeleid is. Scholen en onderwijsvoorrangsgebieden worden door faciliteiten (extra basisformatie en gebiedsfaciliteiten, zie Meijnen & Kloprogge in dit nummer) in staat gesteld te werken aan de verbetering van de onderwijskansen van leerlingen uit achterstandssituaties. De feitelijke uitvoering van het beleid vindt op school- en klasniveau (en op gebiedsniveau) plaats. Uit de formulering van de doelstelling volgt ook dat de beoogde effecten van het beleid gelegen zijn op het niveau van individuele leerlingen. Het beleid is erop gericht de thans bestaande samenhang tussen sociale en etnische achtergrondkenmerken van leerlingen en hun onderwijskansen in basis- en voortgezet onderwijs te verminderen.

Het OVB is dus in essentie een beleid dat gesitueerd is op meerdere niveaus. De consequentie hiervan is dat ook de evaluatie van het OVB op meerdere niveaus dient plaats te vinden en dus een multi-level benadering verdient.

1 Theorie

1.1 Multi-level benadering

De evaluatie van het OVB moet antwoord geven op twee vraagstellingen:

1. wat is het effect van het OVB op de onderwijskansen van de doelgroepen?
2. welke factoren zijn in concreto verantwoordelijk voor de mate waarin het beleid effect sorteert?

Het antwoord op de eerste vraagstelling kan worden verkregen door in een reeks herhaalde metingen na te gaan of de samenhang tussen de sociale en de etnische achtergronden van leerlingen en hun onderwijskansen, uitgedrukt in leerprestaties, kleiner wordt. Aangezien er echter geen sprake is (kan zijn) van een controlegroep van leerlingen, met vergelijkbare achtergrondkenmerken als die van de doelgroepen van het OVB, die niet zijn 'blootgesteld' aan het OVB is echter nooit met zekerheid aan te tonen of de eventuele vooruitgang in prestaties van de doelgroepleerlingen ook daadwerkelijk

het gevolg is van het OVB. Dit probleem wordt nog versterkt door het feit dat de evaluatie pas is gestart op het moment dat het beleid reeds twee jaar in uitvoering was. Wel is het mogelijk om na een reeks van herhaalde metingen bij verschillende cohorten leerlingen vast te stellen of de verschillen tussen de doelgroepleerlingen en een landelijke normgroep (de zogenaamde referentiesteekproef) kleiner worden sedert de invoering van het OVB. Een direct effect van het OVB is echter ook op deze wijze niet vast te stellen.

Wat wel mogelijk is, is nagaan of de verschillende beleidsmiddelen die in het kader van het OVB worden gehanteerd (het formatiebeleid en het gebiedenbeleid) leiden tot verschillende effecten. Daartoe dient gebruik te worden gemaakt van een quasi-experimenteel design waarin leerlingen zijn opgedeeld in groepen die met verschillende beleidsmiddelen in aanraking komen. Deze beleidsmiddelen echter zijn toebedeeld aan scholen en gebieden en niet aan individuele leerlingen. Op dit moment hebben we te maken met een meer-niveau (multi-level) vraagstelling, namelijk de vraag of er verschillen zijn tussen gebieds- en formatiescholen met betrekking tot de prestaties van de doelgroepleerlingen van het OVB. Na adequate controle voor relevante alternatieve verklaringen is het dus in principe mogelijk uitspraken te doen over de effectiviteit van de verschillende OVB-beleidsmiddelen.

Aangezien het niet voor de hand ligt dat er een direct effect van de verschillende beleidsmiddelen zal zijn is in de evaluatie een tweede vraagstelling opgenomen, namelijk de vraag welke factoren in concreto verantwoordelijk zijn voor de mate waarin het beleid - lees de verschillende beleidsmiddelen - effect sorteert (sorteren). Zoals gezegd, is het OVB een faciliteitenbeleid, maar van het verstrekken van faciliteiten aan scholen zijn geen rechtstreekse effecten op leerlingen te verwachten. Scholen en leerkrachten zullen de faciliteiten ook daadwerkelijk moeten inzetten voor activiteiten die gericht zijn op de verbetering van de onderwijskansen van de leerlingen uit achterstandsgroepen. Met andere woorden, het OVB dient op de scholen in activiteiten te worden vertaald om tot enig resultaat te kunnen leiden. Daarom is in het evaluatieprogramma van het OVB naast het 'effectonderzoek' een onderzoek opgenomen naar de manier waarop aan het OVB vorm en inhoud wordt gegeven.

Naar verwachting zal het OVB in de scholen op zeer uiteenlopende wijzen worden geconcretiseerd. Dit maakt het mogelijk na te gaan of een bepaalde invulling meer succesvol is dan een andere of dat sommige invullingen misschien zelfs wel contra-produktief werken. De consequentie hiervan is dat in het onderzoek een breed scala van school- en klaskenmerken moet worden onderzocht die het eventuele effect van de OVB-faciliteiten kunnen mediëren dan wel daarmee interacteren. De algemene vraagstelling van het school- en klaskenmerkenonderzoek van de evaluatie OVB (waarover de rest van dit artikel gaat) is derhalve: via welke school- en klaskenmerken komen de eventuele effecten van het OVB tot stand?

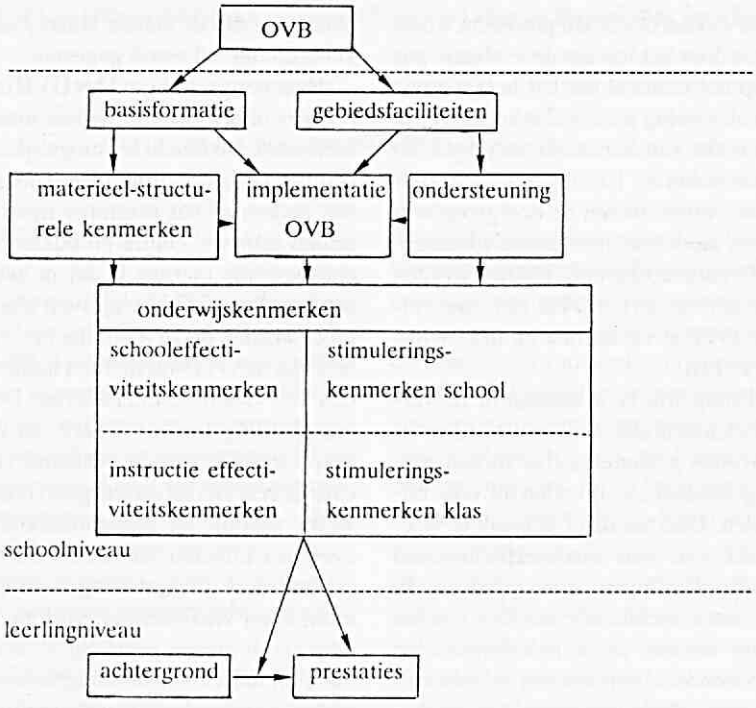
Ook deze vraagstelling is in essentie een multi-level vraagstelling; voor de beantwoording ervan dienen gegevens te worden verzameld bij individuele leerlingen binnen klassen binnen scholen (en in een latere fase binnen gebieden). Om deze reden is een multi-level model ontwikkeld dat kan worden gebruikt voor de verklaring van verschillen tussen scholen in de mate waarin zij erin slagen de onderwijskansen van de doelgroepleerlingen te verbeteren.

1.2 Multi-level onderzoeksmodel en specifieke vraagstellingen

In het ontwikkelde multi-level model zijn twee niveaus onderscheiden, namelijk een schoolniveau en een leerlingniveau. In feite hoort hier tussen in nog het klasniveau, maar aangezien het onderzoek zich afspeelt in het basisonderwijs waar per leerjaar in de meeste gevallen slechts één klas is, valt in de analyse het klasniveau samen met het schoolniveau.

De groepen variabelen die zijn onderscheiden in het bovenste deel van het model zullen hieronder kort worden toegelicht. Het onderste deel van het model (het individuele model) is een sterk vereenvoudigde weergave van het volledige model. Een nadere toelichting hierop wordt gegeven door Van der Werf, Tesser en Mulder (dit nummer).

Bij de keuze van de variabelen op het schoolniveau van het model is er vanuit gegaan dat de vormgeving van het onderwijs het meest essentieel is voor de verbetering van de prestaties van de doelgroepen van het OVB. Om die



Figuur 1. Multi-level model voor de verklaring van prestatieverschillen tussen scholen

reden staat het blok 'onderwijskenmerken' in het model centraal. De onderwijskenmerken zijn onderverdeeld in vier groepen. De groepen 'schooleffectiviteitskenmerken' en 'instructie-effectiviteitskenmerken' zijn gekozen op basis van onderzoek naar kenmerken die een aantoonbaar effect hebben op de leerprestaties van leerlingen in het algemeen, maar waarvan tevens is gebleken dat achterstandsleerlingen daarmee in het bijzonder gebaat zijn. Behalve deze twee groepen kenmerken zijn er echter vanuit de traditie van compensatie- en stimuleringsprogramma's voor achterstandsleerlingen ook andere kenmerken die relevant worden geacht. Hoewel de effectiviteit van deze kenmerken nog niet overtuigend is aangetoond, wordt er in ieder geval op de scholen en in de gebieden veel aandacht aan besteed. Deze kenmerken zijn samengevat onder de noemer 'stimuleringskenmerken' die weer zijn onderscheiden in 'stimuleringskenmerken op schoolniveau' en 'stimuleringskenmerken op klasniveau'.

Voor een optimale vormgeving van het onderwijs op school- en klasniveau met behulp van de faciliteiten die worden verstrekt ten behoeve van achterstandsleerlingen is het nodig dat het OVB ook feitelijk op de scholen een rol speelt, dat wil zeggen dat sprake is van bewust en doelgericht werken aan de uitvoering van het beleid. Met andere woorden: het beleid moet daadwerkelijk in de scholen zijn geïmplementeerd om te kunnen leiden tot een optimale vormgeving van het onderwijs voor achterstandsleerlingen.

Twee andere groepen kenmerken in het model op schoolniveau betreffen de materieel-structurele kenmerken en de mate van ondersteuning van en samenwerking met andere betrokkenen. Deze groepen kenmerken worden beschouwd als randvoorwaarden voor een goede implementatie van het OVB en een goede vormgeving van het onderwijs. Ze kunnen op hun beurt worden bevorderd door de verschillende faciliteitenregelingen waarmee de scholen te maken hebben.

Het model gaat ervan uit dat de verschillende faciliteitenregelingen van het OVB via de school- en klaskenmerken effect hebben op de hoogte van de prestaties van alle leerlingen, gecontroleerd voor de instroomkenmerken van die leerlingen en op de relatie tussen de sociale achtergrondkenmerken van leerlingen en hun prestaties. Of, in statistische termen: het model

veronderstelt dat de schoolvariabelen effect hebben op de parameters van het individuele leerlingmodel. Schooleffecten zijn dan effecten op de intercepten en de slopes van de regressies van prestaties op achtergrondkenmerken binnen scholen.

Op basis van de eerste meting van de evaluatie in het basisonderwijs kan het model reeds op zijn plausibiliteit worden onderzocht. Daartoe zullen in dit artikel de volgende vraagstellingen worden beantwoord:

1 is er een samenhang tussen de omvang en aard van de OVB-faciliteiten enerzijds en de school- en klaskenmerken anderzijds?
2 zijn er verschillen tussen scholen in prestaties van leerlingen (met verschillende achtergrondkenmerken) en zo ja, met welke school- en klaskenmerken hangen deze verschillen samen?

2 Methode

2.1 Opzet van het onderzoek

In het longitudinale onderzoek worden zowel de implementatie van het OVB als de effecten daarvan onderzocht. De onderzoeksopzet kan worden gekarakteriseerd als een combinatie van quasi-experimenteel en longitudinaal onderzoek. Vanwege de organisatorische en financiële haalbaarheid worden geen herhaalde metingen op leerlingniveau uitgevoerd, maar worden scholen of liever gezegd klassen gevolgd. Het gaat hier om de klassen van de leerjaren 4, 6 en 8. Deze klassen worden om de twee jaar getest. Binnen de klassen kan telkens een groot deel van de leerlingen worden teruggevonden, maar ten gevolge van in- en uitstroom van leerlingen kan de klassesamenstelling veranderen. Het quasi-experimentele karakter van het onderzoek is gelegen in het steekproefdesign. De kern van de steekproef is een 3.5% random steekproef uit alle Nederlandse basisscholen. De random steekproef is gestratificeerd op basis van de gebiedenfactor en de schoolscore factor. De gebiedenfactor heeft twee niveaus, scholen die tot een gebied behoren en niet-gebiedsscholen. De schoolscorefactor heeft drie niveaus, die de sociale samenstelling van de leerlingpopulatie van de school weerspiegelen. Deze drie niveaus zijn scholen met een schoolscore kleiner dan 105, scholen met een score tussen 105 en 115 en scholen met een score hoger dan 115. Deze stratificatie van de kernsteekproef resulteert in zes strata. Strata met minder dan 100 scholen zijn aangevuld tot 100. De strata met gebiedsscholen zijn aangevuld met alle scholen uit de helft van de 70 gebieden, zodat ook analyses op gebiedsniveau kunnen worden uitgevoerd. Het stratum 'gebiedscholen met een schoolscore kleiner dan 105' is leeg omdat deze scholen niet bestaan.

De implementatie van het beleid is onderzocht op dezelfde scholen als die waarop de geteste leerlingen zitten. Daarvoor zijn gegevens verzameld over school- en klaskenmerken. De gegevens over de implementatie van het beleid kunnen worden gerelateerd aan de prestatiescores, zodat factoren kunnen worden opgespoord waaraan de eventuele effecten van het beleid kunnen worden toegeschreven.

2.2 Variabelen en instrumenten

In het onderzoek zijn twee hoofdgroepen van variabelen onderscheiden, te weten variabelen op school- en klasniveau en variabelen op leerlingniveau.

De variabelen op school- en klasniveau zijn alle gemeten met behulp van schriftelijke vragenlijsten die moesten worden ingevuld door de directeuren en door de leerkrachten die les gaven aan de geteste groepen 4, 6 en 8. Informatie over de inhoud van de vragenlijsten wordt gegeven bij de bespreking van de resultaten. Voor de psychometrische gegevens wordt verwezen naar Van der Werf, Weide en Tesser (1991). De variabelen op leerlingniveau zijn in de eerste plaats de algemene kenmerken leerlinggewicht, geslacht, geboortedatum en etnische herkomst. Informatie hierover is verkregen van de leerlingen zelf en van de leerkracht of de directeur van de school. In de tweede plaats zijn de prestaties van de leerlingen gemeten met behulp van toetsen voor taal, rekenen en intelligentie. In de derde plaats zijn gegevens verzameld over de opleiding en het beroep van de ouders, over hun participatie aan culturele activiteiten en over de mate waarin zij onderwijsondersteunend gedrag vertonen. Al deze gegevens zijn verzameld met behulp van schriftelijke vragenlijsten voor de ouders.

2.3 Procedure

In mei 1988 zijn de scholen uit de steekproef schriftelijk benaderd met informatie over de evaluatie en met het verzoek om medewerking daaraan te verlenen. In juni/juli zijn de scholen gebeld. Aan de scholen die wilden meewerken is gevraagd gegevens te verstrekken over aantallen leerlingen, combinatieklassen, parallelklassen e.d.

In augustus zijn de scholen die bereid waren mee te werken gebeld voor het maken van afspraken door de testleiders die deze scholen gingen bezoeken. De testleiders hadden van te voren een uitgebreide instructie en training voor de testafname gehad. Elke school is in principe drie dagen bezocht voor het afnemen van de toetsen en voor het inzamelen van de directie- en leerkrachtvragenlijsten.

2.4 Respons

Uitgaande van de gekozen steekproefopzet zouden in totaal 816 scholen moeten meedoen. Om aan dit aantal te komen zijn op basis van de verwachte non-respons ongeveer 1250 scholen benaderd met het verzoek om medewerking. Hiervan bleken 696 scholen (=56%) bereid aan het onderzoek mee te doen. Dit aantal is beduidend lager dan de 816 scholen uit de steekproefopzet. De oorzaak hiervan is de onverwachte non-respons van de gebiedsscholen. Ondanks deze hogere uitval dan verwacht is de representativiteit van de referentiegroep goed. De representativiteit van de groep gebiedsscholen is iets minder; gebiedsscholen in de grote steden zijn ondervertegenwoordigd, die in de middelgrote steden en plattelandsgemeenten oververtegenwoordigd.

Binnen de deelnemende scholen hebben, op een enkele uitzondering na, alle leerlingen uit alle groepen 4, 6 en 8 meegedaan aan de testafnames. Van de 696 scholen hebben 641 directeuren (=93%) de vragenlijst ingevuld. De respons van de directeuren van de gebiedsscholen met een schoolscore hoger dan 115 is wat lager (89%) dan die van de directeuren van de andere groepen scholen (93-98%). Het aantal groepsleerkrachten dat een vragenlijst heeft ingevuld bedraagt (inclusief directeuren die zelf ook groepsleerkracht zijn) ruim 2000, dus gemiddeld ongeveer drie groepsleerkrachten per school. Voor details over de respons en de representativiteit verwijzen we naar Tesser, Mulder en Van der Werf (1991).

3 Resultaten

3.1 Verschillen tussen scholen in school- en klaskenmerken

Ter beantwoording van de eerste onderzoeksvraag of er een samenhang is tussen de omvang en de aard van de OVB-faciliteiten enerzijds en de school- en klaskenmerken anderzijds is voor alle school- en klaskenmerken geanalyseerd of er significante verschillen zijn tussen de categorieën scholen die in verschillende mate en op verschillende wijze OVB-faciliteiten ontvangen. De volgende vijf categorieën scholen zijn onderscheiden:

A scholen zonder onderwijsvoorrangsfaciliteiten (schoolscore < 105);
B gebiedscholen met een redelijke hoeveelheid extra basisformatie (schoolscore 105-115);
C niet-gebiedsscholen met een redelijke hoeveelheid extra basisformatie (schoolscore tussen 105 en 115);
D gebiedsscholen met veel extra basisformatie (schoolscore > 115);
E niet-gebiedsscholen met veel extra basisformatie (schoolscore > 115).

Verondersteld wordt dat naarmate scholen meer faciliteiten hebben meer sprake zal zijn van de voor OVB relevant geachte kenmerken. Dus categorie E zal hoger scoren dan categorie D, deze weer hoger dan categorie C, en deze weer hoger dan categorie B. Categorie A zal het laagst scoren op de kenmerken.

De verschillen zijn getoetst met *F*-toetsen. Een verschil is alleen als relevant beschouwd als $p<.001$. In de appendix wordt een overzicht van de resultaten gegeven. Hieronder volgt een samenvatting.

Met betrekking tot de materieel-structurele kenmerken blijkt dat de scholen in groep D aanmerkelijk meer formatie hebben en tevens veel meer kinderen uit etnische minderheden. De werkdruk op deze scholen is iets hoger dan die op de andere groepen scholen.

De implementatie van het beleid is beter op de scholen die deelnemen aan een onderwijsvoorrangsgebied (groep B en D) dan in de andere groepen scholen. De attituden ten opzichte van het OVB van de onderwijsgevenden zijn positiever; ze participeren vaker actief in voorrangsactiviteiten; ze hebben meer kennis van

en ervaring met kinderen uit achterstandssituaties. Gebiedsscholen krijgen tevens meer ondersteuning van ondersteunings- en welzijnsinstellingen.

De verschillen tussen de groepen scholen in schooleffectiviteitskenmerken zijn erg klein. Gebiedsscholen (groep B en D) scoren iets hoger op 'teamcohesie' en iets lager op 'aspiratieniveau'.

Met betrekking tot de stimuleringskenmerken op schoolniveau blijkt dat de scholen met veel allochtone leerlingen (groep D en E) vaker aandacht besteden aan intercultureel onderwijs en onderwijs in Eigen Taal en Cultuur (OETC). Verder onderhouden de scholen in groep D meer contacten met de ouders en de leerlingen.

De verschillen in instructie effectiviteitskenmerken zijn eveneens zeer klein. In groep D geven leerkrachten wat meer klassikaal les, geven vaker huiswerk mee en besteden meer lestijd aan taal en rekenen.

Ten slotte blijkt met betrekking tot de stimuleringskenmerken op klasniveau dat op de scholen in groep D leerkrachten vaker gebruik maken van specifieke materialen voor achterstandsleerlingen, meer accent leggen op communicatief taalgebruik, meer doen aan taalstimulering en minder contact hebben met ouders.

Samenvattend, er zijn niet veel verschillen tussen de groepen scholen. Als er al verschillen zijn, zijn ze erg klein. De gebiedsscholen met veel extra basisformatie wijken dan meestal af van de andere groepen scholen. Al met al lijkt er nauwelijks sprake van samenhang tussen de aard en omvang van de OVB-faciliteiten en de school- en klaskenmerken.

3.2 Effecten van school- en klaskenmerken op de leerlingprestaties

Ter beantwoording van de tweede onderzoeksvraag of er verschillen zijn tussen scholen in prestaties van leerlingen (met verschillende achtergrondkenmerken) en met welke school- en klaskenmerken de eventuele verschillen samenhangen zijn multi-level analyses uitgevoerd met het multi-level programma VARCL (Longford, 1988). De gegevens betreffen in principe vier niveaus, namelijk het leerling-, school-, klas- en gebiedsniveau. Omdat er echter geen gegevens op gebiedsniveau verzameld zijn, maar alleen de directeuren van de scholen zijn bevraagd over een aantal gebiedskenmerken, beschouwen we de gebiedsgegevens als kenmerken op schoolniveau. Er blijven dus drie analyseniveaus over. Concreet gaat het om leerlingen in groep 4, 6 en 8 van zo'n 700 basisscholen. Bij de leerlingen van deze groepen zijn groepsspecifieke toetsen afgenomen, hetgeen betekent dat de drie groepen niet gezamenlijk geanalyseerd kunnen worden. Wanneer de analyses per groep uitgevoerd worden, vallen school- en klasniveau in feite samen, omdat op verreweg de meeste scholen per leerjaar slechts één leerkracht met een groep is. Daarom is besloten de schoolkenmerken voor de analyses te disaggregeren naar het klasniveau, zodat uiteindelijk slechts twee analyseniveaus overblijven: de leerling en de docent.

In de analyses zijn als afhankelijke variabelen op leerlingniveau opgenomen de taal- en rekenscores en als covariaten intelligentie, weegfactor en geslacht. Voor de eerste selectie van de school/docentvariabelen als onafhankelijke variabelen is uitgegaan van de gevonden verschillen tussen groepen scholen. Deze selectie resulteerde in een te grote hoeveelheid variabelen voor de VARCL-analyses. Daarom zijn voor de uiteindelijke selectie Pearson correlatiecoëfficiënten berekend tussen deze variabelen en de taal- en rekenscores. Alleen indien de correlatie hoger was dan .15 werd de variabele in de VARCL-analyses opgenomen. Na deze selectie bleven 19 variabelen over. Deze zijn in de appendix aangeduid met een sterretje.

Allereerst is een zogenaamd 'leeg' model gespecificeerd, dat wil zeggen een model zonder leerling co-variaten en zonder school- en klaskenmerken. Hierdoor worden de proporties te verklaren variantie op leerling- en schoolniveau verkregen.

In het 'lege model', is te zien dat het percentage variantie op schoolniveau voor beide vakken relatief hoog is (zie Tabel 1) in vergelijking met de resultaten van onderzoek van onder andere Brandsma en Knuver, 1989; Reezigt en Weide, 1989.

Er zijn grotere verschillen tussen scholen in taalprestaties dan in rekenprestaties. Dit is niet erg gebruikelijk. Een verklaring hiervoor moet waarschijnlijk gezocht worden in het feit dat het hier niet om een a-selecte steekproef gaat,

Tabel 1
Proporties variantie op leerling- en schoolniveau, per groep, per vak ('leeg' model)

	TAAL			REKENEN		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
leerlingniveau	70%	80%	79%	74%	82%	77%
schoolniveau	30%	20%	21%	26%	18%	23%
totaal	100%	100%	100%	100%	100%	100%

Tabel 2
Proporties varianties op leerling- en schoolniveau, per groep, per vak (model met leerlingcovariaten)

	TAAL			REKENEN		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
leerlingniveau	59%	66%	63%	59%	62%	55%
schoolniveau	16%	7%	6%	15%	8%	11%
totaal	75%	73%	69%	74%	70%	66%
verklaarde variantie	25%	27%	31%	26%	30%	34%

maar om een oververtegenwoordiging van scholen met relatief veel leerlingen uit laag sociaal milieu en scholen met relatief veel allochtone leerlingen.

De covariaten intelligentie, gewicht en sekse hebben in bijna alle gevallen een significante relatie met de prestaties (bij eenzijdige toetsing op 1%-niveau). Slechts in één geval is dit niet zo: in groep 8 is bij taal geen relatie gevonden tussen sekse en prestaties. De covariaten verklaren samen bijna 30% van de totale variantie (zie Tabel 2).

In de meeste gevallen wordt van de oorspronkelijke verschillen tussen scholen in het lege model meer dan de helft verklaard door de intelligentie, de weegfactor en de sekse van de individuele leerlingen op de scholen gezamenlijk. Bij taal neemt de weegfactor hiervan telkens het grootste deel voor zijn rekening, bij rekenen is dit intelligentie. Dit ondersteunt het eerder geuite vermoeden dat de grote verschillen in leerlingpopulatie vooral verantwoordelijk zijn voor de oorspronkelijk grote verschillen tussen scholen bij taal en de wat geringer, maar toch nog grote verschillen tussen scholen bij rekenen. In groep 4 wordt het minst van de tussen-scholen variantie verklaard. De verschillen tussen scholen die overblijven na correctie voor de individuele leerlingkenmerken zijn nu, met uitzondering van groep 4, groter bij rekenen dan bij taal en komen nu vrij nauwkeurig overeen met resultaten van ander onderzoek in het basisonderwijs (Brandsma & Knuver, 1989; Tesser e.a., 1991). Het zijn deze verschillen die in principe verklaard kunnen worden door de school- en klaskenmerken.

Vervolgens zijn alle geselecteerde school- en klaskenmerken in het model opgenomen (volledige model). De kenmerken die niet significant bijdragen aan de verklaring van de verschillen tussen scholen (bij eenzijdige toetsing op 5% niveau) zijn vervolgens weer verwijderd. Uit de resultaten blijkt dat de verschillen tussen scholen slechts in geringe mate worden verklaard door de school- en klaskenmerken (zie Tabel 3).

Inspectie van de proporties onverklaarde

Tabel 3
Proporties variantie op leerling- en schoolniveau, per groep, per vak (model met leerlingcovariaten en school- en klaskenmerken)

	TAAL			REKENEN		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
leerlingniveau	59%	66%	63%	59%	62%	55%
schoolniveau	13%	6%	5%	12%	6%	9%
totaal	72%	72%	68%	71%	68%	64%
verklaarde variantie	28%	28%	32%	29%	32%	36%

variantie die na opname van de school- en klaskenmerken resteren, laat zien dat deze kenmerken in totaal tussen de 14% en 25% van de resterende tussen-scholen variantie, wanneer deze op 100% wordt gesteld, verklaren. Van de totale variantie verklaren de school- en klaskenmerken tussen 1% en 3%.
Verder blijkt dat er per groep en per vak nogal wat verschillen zijn in de kenmerken die er iets toe lijken te doen. Daarop wordt verderop ingegaan.

Tot nu toe is alleen gekeken naar verschillen in de gemiddelde prestaties tussen scholen, gecorrigeerd voor de achtergrondkenmerken van individuele leerlingen. De vraag of er verschillen zijn tussen scholen in de prestaties van leerlingen (met verschillende achtergrondkenmerken) en met welke schoolkenmerken deze samenhangen is hiermee slechts gedeeltelijk beantwoord. Nog onbeantwoord is het deel van de vraag dat betrekking heeft op verschillen tussen scholen in prestaties van leerlingen met verschillende achtergrondkenmerken. Het betreft hier de vraag naar de samenhang tussen achtergrondkenmerken van leerlingen en prestaties. Het is mogelijk dat deze relatie niet op alle scholen hetzelfde is, maar dat sommige scholen betere prestaties weten te bereiken bij bijvoorbeeld allochtone kinderen dan andere scholen, terwijl de prestaties van 1.00 kinderen op beide groepen scholen gelijk zijn. De eerstgenoemde groep scholen heeft dan een beter compenserend vermogen dan de laatstgenoemde (vgl. Brandsma & Knuver, 1988). Indien hiervan sprake is moeten de regressielijnen van de regressie van prestaties op leerlingkenmerken van individuele scholen afwijken van een 'gemiddelde' regressielijn die de algemene relatie tussen een leerlingkenmerk en de prestaties weergeeft. De variantie van de regressielijnen wordt hellingvariantie genoemd.

Om de verschillen in compenserend vermogen tussen scholen te onderzoeken zijn in het multi-level model de leerlingvariaten intelligentie, gewicht en sekse enerzijds en prestaties anderzijds ook nog eens 'variabel' gemodelleerd. Indien sprake is van een beter 'modelfit' mag worden gesproken over verschillen in compenserend vermogen tussen scholen. Uit de analyses blijkt dat hiervan nauwelijks sprake is. Alleen bij taal in groep 8 en bij rekenen in groep 4 is er een verschil tussen scholen in de relatie tussen intelligentie en prestaties en tussen weegfactor en prestaties. Dit betekent dat sommige scholen er beter in slagen voor de verschillen tussen leerlingen die samenhangen met intelligentie en de weegfactor te compenseren. Wat betreft het compenserend vermogen voor intelligentie zijn de verschillen vrij klein. De hellingvariantie bedraagt .01 bij taal in groep 8 en .02 bij rekenen in groep 4. Het verschil tussen scholen in compenserend vermogen voor weegfactor is veel groter maar betreft alleen de 1.90 en 1.00 leerlingen: op sommige scholen is het verschil tussen deze beide groepen leerlingen minder groot dan op andere. De hellingvarianties bedragen .10 voor rekenen in groep 4 en .16 voor taal in groep 8. Concreet betekent een hellingvariantie van .16 bij taal in groep 8 voor de 1.90 leerlingen op de slechtst compenserende school dat zij een verwachte taalscore hebben op een toets van 67 items die ongeveer 12.9 punten lager ligt dan die van de 1.00 leerlingen, terwijl dit verschil op de best compenserende school slechts .8 punten is. Helaas zijn er geen school- en klaskenmerken die de verschillen in compenserend vermogend tussen scholen kunnen verklaren.

4 Interpretatie van de resultaten van de multilevel analyse

4.1 Leerlingkenmerken

Om inzicht te krijgen in de betekenis van de gevonden relaties tussen de leerlingcovariaten intelligentie, gewicht en sekse worden in Tabel 4 de regressiecoëfficiënten voor de regressies van prestaties op de leerlingkenmerken gepresenteerd.

Intelligentie hangt, zoals verwacht, positief samen met prestaties. De weegfactor hangt telkens negatief samen, dat wil zeggen dat de 1.25- en 1.90 leerlingen gemiddeld lagere taal- en rekenprestaties hebben dan de 1.00 leerlingen. Ook dit is overeenkomstig de verwachting; het gaat immers om leerlingen die juist een gewicht hebben gekregen vanwege de te verwachten nadelige effecten van hun sociaal-economische achtergrond. Ten slotte blijkt dat meisjes het beter doen bij taal, behalve in groep 8, en bij rekenen een duidelijke achterstand hebben in vergelijking met jongens.

Tabel 4

Regressiecoëfficiënten voor de regressie van prestaties op leerlingkenmerken, per groep, per vak (model met alleen leerlingcovariaten)

	TAAL			REKENEN		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
intelligentie	.27	.32	.34	.39	.44	.45
gewicht 1.25	-.20	-.34	-.28	-.16	-.26	-.27
gewicht 1.90	-.72	-.80	-.91	-.41	-.50	-.57
sekse (meisje)	.08	.12	–	-.11	-.20	-.23

Vergelijking van de regressiecoëfficiënten onderling laat zien dat bij taal in alle groepen de (negatieve) invloed van een gewicht van 1.90 verreweg het grootst is. Een intelligente 1.90 leerling zal dus een verwachte taalscore hebben die lager is dan de verwachte taalscore van een minder intelligente 1.00 leerling. De invloed van intelligentie en een gewicht van 1.25 is bij taal ongeveer even groot. Bij rekenen is de invloed van gewicht vergeleken met die van intelligentie wat kleiner, hoewel ook hier geldt dat de invloed van gewicht telkens het grootst is. In alle gevallen doet de sekse van de leerling er het minst toe. Wel valt bij rekenen op dat sekse en milieu (gewicht 1.25) bijna evenveel invloed hebben op de verwachte prestaties.

Omdat de prestatiescores gestandaardiseerd zijn, zijn de regressiecoëfficiënten in de verschillende VARCL-analyses vergelijkbaar. Vergelijking van de groepen laat zien dat de invloed van de covariaten in groep 4 steeds het kleinst is, al zijn de verschillen tussen de groepen niet erg groot.

De concrete betekenis van de regressiecoëfficiënten kan worden bepaald door twee leerlingen op basis van de door de VARCL berekende schattingen van de regressiecoëfficiënten te vergelijken. De ene leerling heeft een gewicht van 1.00, de ander van 1.90. Ze hebben dezelfde intelligentiescore en zijn van dezelfde sekse. Wanneer de leerlingen in groep 8 zitten zal de 1.90 leerling naar verwachting een taalscore behalen die negentiende standaarddeviatie lager ligt dan de score van de 1.00 leerling (de regressiecoëfficiënt is -.91). Op de taaltoets van 67 items gaat het concreet om een score die 5.8 punten lager ligt. De rekenscore van de-

Tabel 5

Regressiecoëfficiënten voor de regressie van prestaties op school- en klaskenmerken (model met leerlingcovariaten en significante school- en klaskenmerken)

	TAAL			REKENEN		
	groep 4	groep 6	groep 8	groep 4	groep 6	groep 8
– aantal leerlingen per formatieplaats	.71			1.28		.55
– % allochtone leerlingen in de groep	-.27	-.39			-.31	
– beleid gericht op achterstandsleerlingen			-.08		-.05	
– op de hoogte OVB			.05			
– deelname OVB-activiteiten		-.07	-.07			
– oorzaken achterstand bij school leggen				-.07	-.06	
– kennis van/ervaring met onderwijsaanpak						
– achterstandsleerlingen	.07					
– verwachte uitstroom naar IBO/LBO	-.08					
– wijze van ICO	-.07					
– klassikaal werken				.07	.05	
– minimumdoelen taal			.07			
– minimumdoelen rekenen			-.11			-.11
– huiswerk			.05	.07	.06	.07

zelfde 1.90 leerling zal ongeveer zestiende standaarddeviatie (de regressiecoëfficiënt is -.57) lager liggen dan de rekenscore van de 1.00 leerling, omgerekend zo'n 4.5 punten op een rekenscore van 40 items. Dit zijn aanzienlijke verschillen. Voor een 1.25 leerling met eenzelfde intelligentie en van dezelfde sekse in groep 8 is het verschil met de 1.00 leerling wat kleiner, maar toch nog ongeveer drietiende standaarddeviatie. Bij taal gaat het concreet om 2.3 punten, bij rekenen om 2.2 punten. Hierbij vergeleken zijn de verschillen tussen jongens en meisjes relatief klein. Meisjes doen het bij taal ongeveer eentiende standaarddeviatie beter dan jongens (behalve in groep 8 waar geen verschil is) en bij rekenen tweetiende standaarddeviatie (in groep 4 eentiende standaarddeviatie) lager.

4.2 School- en klaskenmerken

In Tabel 5 staan voor de school- en klaskenmerken die in de VARCL-analyses significant bleken samen te hangen met de prestaties de regressiecoëfficiënten voor de regressies van prestaties op deze kenmerken.

Bij het interpreteren van de regressiecoëfficiënten moet ermee rekening gehouden worden dat de variabelen 'aantal leerlingen per formatieplaats' en 'percentage allochtone leerlingen' niet zijn gestandaardiseerd. Hierdoor zijn de regressiecoëfficiënten van deze kenmerken niet goed vergelijkbaar met die van de overige kenmerken.

Beschouwing van de materieel-structurele kenmerken laat zien dat het feit of een school een gebiedsschool is of niet er in geen enkele analyse iets toe doet. Dit is op zich niet zo verwonderlijk omdat dit kenmerk overlapt met allerlei andere kenmerken zoals het aantal leerlingen per formatieplaats, het percentage allochtone leerlingen en implementatiekenmerken. In drie analyses is een positieve samenhang gevonden tussen het aantal leerlingen per formatieplaats en prestaties, een keer bij taal en twee keer bij rekenen. Dus hoe minder een school aan formatie heeft, des te hoger zijn de gemiddelde prestaties. Evenzo wordt drie keer een negatief verband gevonden tussen het percentage allochtone leerlingen in de groep en prestaties, ditmaal twee keer bij taal en een keer bij rekenen.

Van de implementatiekenmerken vertonen het voeren van een beleid gericht op achterstandsleerlingen en de mate waarin men deelneemt aan OVB-activiteiten in twee gevallen een negatieve samenhang met prestaties. Op de hoogte zijn van OVB hangt een keer positief samen met prestaties. Ondersteuning van de onderwijsbegeleidingsdienst speciaal gericht op achterstandsleerlingen en de frequentie waarmee contacten met welzijnsinstellingen worden onderhouden worden in de tabel niet teruggevonden. Deze kenmerken hangen dus in geen enkel geval samen met leerprestaties van leerlingen. De mate waarin leerkrachten de oorzaken van onderwijsachterstanden bij de school leggen hangt twee keer negatief samen met rekenprestaties. Kennis van en ervaring met een onderwijsaanpak gericht op achterstandsleerlingen hangt een keer positief samen met taalprestaties.

Van het blok 'effectieve instructiekenmerken' hangt de verwachte uitstroom naar IBO/LBO een keer negatief samen met taalprestaties. Dit moet worden geïnterpreteerd als 'naarmate men van meer leerlingen verwacht dat ze uitstromen naar een laag type voortgezet onderwijs', de taalprestaties lager zijn. Opvallend is dat dit juist in groep 4 het geval is. Prestatiegerichtheid doet er in geen enkele analyse iets toe. Verder springt het kenmerk huiswerk er in positieve zin uit. Er is vier keer een positieve samenhang gevonden tussen de frequentie waarmee leerkrachten huiswerk geven en prestaties, bij rekenen in alle groepen en bij taal in groep 8. Daarnaast vertonen minimumdoelen en klassikaal werken in twee gevallen een positieve samenhang met prestaties. De negatieve regressiecoëfficiënt van rekenen moet zo worden geïnterpreteerd dat leerkrachten die voor alle leerlingen dezelfde rekendoelen nastreven leerlingen in de klas hebben die gemiddeld hoog presteren. Leerkrachten die bij rekenen door individuele leerlingen verschillende doelen nastreven hebben daarentegen leerlingen in de klas die gemiddeld laag presteren. Vreemd genoeg hangt het nastreven van minimumdoelen bij taal in groep 8 negatief samen met de taalprestaties in die groep (de regressiecoëfficiënt is positief). Klassikaal werken hangt alleen bij rekenen positief samen met prestaties, bij taal is geen samenhang gevonden.

Ten slotte blijkt in het blok 'stimuleringskenmerken op klasniveau' de vorm van inter-

cultureel onderwijs een keer en wel negatief samen te hangen met de prestaties. De frequentie van intercultureel onderwijs doet er in geen enkele analyse iets toe, evenmin als ouderparticipatie. Samenvattend moet geconcludeerd worden dat er bepaald geen sprake is van een eenduidig beeld. Alleen het kenmerk 'huiswerk' hangt redelijk consistent samen met prestaties. De regressiecoëfficiënten van dat kenmerk liggen tussen .05 en .07. Concreet betekent dit bijvoorbeeld voor de taalprestaties in groep 8 dat leerlingen op een school waar een standaarddeviatie meer dan gemiddeld huiswerk wordt gegeven, ongeveer een twintigste standaarddeviatie (regressiecoëfficiënt van .05) hoger presteren dan op een school waar een gemiddelde hoeveelheid huiswerk wordt gegeven.

5 Conclusies

De conclusies uit de resultaten die hierboven beschreven zijn, zijn tweeledig. In de eerste plaats moet worden geconstateerd dat er niet veel opvallende verschillen zijn tussen scholen in de wijze waarop ze hun onderwijs op school- en klasniveau organiseren en de wijze waarop ze lesgeven aan en werken met leerlingen uit achterstandssituaties. Als er al verschillen zijn is het bijna altijd de groep gebiedsscholen met een hoge schoolscore die afwijkt van de andere groepen. Dit is ook de groep scholen die het meest te maken heeft met een hoge concentratie allochtone leerlingen (gemiddeld 50%). Noch de extra basisformatie noch de gebiedencomponent differentieert tussen scholen voor wat betreft de school- en klaskenmerken. Dit wijst erop dat de mate waarin leerkrachten aandacht hebben voor effectieve school- en instructiekenmerken maar weinig te maken heeft met de aard en omvang van de OVB-faciliteiten.

Deze conclusie wordt ondersteund door de bevindingen ten aanzien van de relaties tussen school- en klaskenmerken enerzijds en de prestaties van leerlingen anderzijds. Uit de multilevel analyses komt naar voren dat de variantie tussen scholen na controle voor de achtergrondkenmerken van de leerlingen erg klein is. Opvallend is dat deze variantie niet in alle leerjaren gelijk is. In groep 4 is de variantie duidelijk groter dan in groep 6 en 8, zowel voor taal als voor rekenen. In groep 6 is de variantie voor taal iets hoger dan in groep 8, voor rekenen geldt het omgekeerde. Een duidelijke verklaring voor deze verschillen in varianties over de leerjaren heen valt voorlopig niet te geven. Mogelijk hebben ze iets te maken met de stabiliteit van schooleffecten over jaargroepen heen (vgl. Bosker, 1991). In de analyse van de tweede meting zal worden geanalyseerd hoe het met deze stabiliteit zit.

Uit de multi-level analyse komt verder naar voren dat de voorrangsfaciliteiten niet relevant zijn voor de verklaring van de variantie tussen scholen; de gebiedsfaciliteiten verklaren niets, de formatiecomponent draagt in negatieve zin bij. Het maakt dus voor een leerling, ongeacht zijn of haar sociale of etnische achtergrond, geen verschil of hij of zij op een gebiedsschool zit of niet.

Gebiedsscholen zijn tevens niet in staat beter te compenseren voor de sociale of etnische afkomst van leerlingen dan niet-gebiedsscholen. De verschillen tussen 1.25- en 1.90-leerlingen enerzijds en 1.00-leerlingen anderzijds zijn op gebiedsscholen even groot als op niet-gebiedsscholen.

De extra basisformatie blijkt een negatief effect te hebben op de prestaties van leerlingen, na controle voor de achtergrondkenmerken van de leerlingen. Hoe meer formatie een school heeft, des te lager zijn de prestaties. De hoeveelheid formatie hangt samen met de leerlingpopulatie van een school. Dit betekent dat het voor individuele leerlingen, ongeacht hun sociale achtergrond, wel degelijk uitmaakt of ze een school bezoeken waarop veel achterstandsleerlingen zitten of een school met maar weinig van deze leerlingen. Voor de verschillen tussen scholen wat betreft de samenhang tussen achtergrondkenmerken en prestaties maakt de hoeveelheid extra basisformatie, i.c. de samenstelling van de leerlingpopulatie, echter niet uit. Leerlingen met een gewicht van 1.90 hebben ten opzichte van 1.00-leerlingen een even grote achterstand op scholen waar veel achterstandsleerlingen zitten als op scholen waar maar weinig van die leerlingen zitten. Dit resultaat is in strijd met hetgeen Tesser e.a. (1991) eerder hebben gevonden voor adviezen voor het voortgezet onderwijs.

Voor het overige is het beeld van de samenhang tussen school- en klaskenmerken en de

prestaties van leerlingen tamelijk diffuus. De kenmerken verschillen al naar gelang het vak en het leerjaar waar het om gaat. Het enige consistente resultaat is de frequentie waarmee leerkrachten huiswerk geven. In vier van de zes analyses heeft deze variabele een positief effect. Dat betekent dat elke willekeurige leerling, ongeacht zijn of haar sociale achtergrond, beter zal presteren op een school waar meer huiswerk wordt gegeven. Echter, de verschillen tussen scholen in de relatie tussen sociale achtergrond en prestaties kunnen niet door de variabele huiswerk worden verklaard. Het negatieve effect van de sociale achtergrond lijkt dus niet te kunnen worden gecompenseerd door de leerlingen vaker huiswerk mee te geven.

Ten slotte moet geconstateerd worden dat de in dit onderzoek gespecificeerde school- en klaskenmerken kennelijk niet goed in staat zijn het belangrijkste deel van de verschillen tussen scholen te verklaren. Vooral in groep 4 blijkt nog veel onverklaarde tussen-scholen variantie over. Dit betekent dat bepaalde school- en klaskenmerken die voor het verklaren van deze verschillen van belang zijn, in het model ontbreken. Dit geldt in het bijzonder voor de verschillen tussen scholen in compenserend vermogen. Op zich is het echter voor het OVB bemoedigend dat deze verschillen er zijn, omdat ze bewijzen dat sommige scholen wel degelijk in staat zijn te compenseren voor de nadelige effecten van de sociale en etnische achtergrond van leerlingen. Het opsporen van variabelen die hiervoor verantwoordelijk zijn, lijkt op dit moment nog slechts een kwestie van tijd en het vinden van de juiste meetprocedures.

Literatuur

Bosker, R. R. (1991). De consistentie van schooleffecten in het basisonderwijs, *Tijdschrift voor Onderwijsresearch, 16*, 206-219.

Brandsma, H. P., & Knuver, J. W. M. (1988). Organisatorische verschillen tussen basisscholen en hun effect op leerlingprestaties. *Tijdschrift voor Onderwijsresearch, 13*, 201-212.

Brandsma, H. P., & Knuver, J. W. M. (1989). *Basisschoolkenmerken als determinanten van het functioneren van leerlingen.* Deelrapport II, Resultaten. Groningen: RION.

Longford, N. (1988). *VARCL manual.* Princeton, NJ: Educational Testing Service.

Ministerie van Onderwijs en Wetenschappen (1985). *Onderwijsvoorrangsplan.* 's Gravenhage: Staatsuitgeverij.

Reezigt, G. J., & Weide, M. G. (1989). *Effecten van differentiatie. Resultaten survey-onderzoek* (deelrapport IV). Groningen: RION.

Tesser, P., Mulder, L., & Werf, G. van der. (1991). *De eerste fase van de longitudinale OVB onderzoeken. Het leerlingenonderzoek.* Nijmegen/Groningen: RION/ITS.

Werf, M. P. C. van der, Weide, M., & Tesser, P. (1991). *Het onderwijsvoorrangsbeleid in de school en in de klas. De eerste meting in het basisonderwijs.* Groningen/Nijmegen: RION/ITS.

Manuscript aanvaard 11-1-1993

Auteurs

M. P. C. van der Werf is werkzaam bij het RION, Instituut voor Onderwijsonderzoek van de Rijksuniversiteit Groningen als projectleider van de landelijke evaluatie van het OVB.

M. G. Weide was tot medio 1992 als assistent in opleiding op het RION, Instituut voor Onderzoek van het Onderwijs van de Rijksuniversiteit Groningen, werkzaam bij de evaluatie van het OVB. Momenteel werkt zij bij het Centrum Onderzoek van het Wetenschappelijk Onderwijs Groningen (COWOG).

Correspondentie-adres: RION, Instituut voor Onderwijsonderzoek, Postbus 1286, 9701 BG Groningen.

Abstract

Effective priority characteristics in the school and in the classroom

M. P. C. van der Werf & M. G. Weide. Pedagogische Studiën, 1993, 70, 108-121.

In the first cohort study of the national evaluation of the Dutch Educational Priority Policy, carried out in 1988, information about school and classroom characteristics is gathered by questionnaires to be filled in by principals and teachers. Students in grade 4, 6 and 8 were tested for intelligence, arithmetic and

Dutch language.

From the results it appears that there are some differences in school and classroom characteristics between schools with different amounts of disadvantaged children, but these differences are not very large. Schools with a lot of ethnic minority children score a bit higher on some school and instructional effectiveness variables and on some specific strategies for disadvantaged children compared to the other groups of schools. However, from the multilevel analyses it appears that almost none of these variables is related with differences between schools in students' achievements. Only the frequency of homework shows a rather consistent relationship with arithmetics achievement.

Appendix

Tabel

*School- en klaskenmerken; verschillen tussen groepen scholen (p <.001); variabelen in VARCL-analyses**

Variabele	*Verschillen tussen cellen*	*VARCL*
Materieel-structurele kenmerken		
Gebiedsfaciliteiten	Groep B, D >	*
Basisformatie	Groep D >	*
Schoolgrootte	Groep A, D >	
% etnische minderheden	Groep D >	*
Ervaring van leerkrachten	Groep B >	
Werkklimaat	Groep D >	
Implementatiekenmerken		
Aandacht OVB	Groep B, D >	*
Attitude t.o.v. OVB	Groep B, D >	
Kennis van OVB	Groep B, D >	*
OVB-activiteiten	Groep B, D >	*
Oorzaken achterstanden bij school	Groep D <	*
Ervaring met achterstanden	Groep D >	*
Samenwerking met andere scholen	Groep D >	
Ondersteuning	Groep D >	*
Samenwerking met welzijnsinstellingen	Groep D >	*
Schooleffectiviteitskenmerken		
Ordelijk schoolklimaat	n.s.	
Nadruk op regels	n.s.	
Verwachte uitstroom IBO/LBO	Groep B, D >	*
Registratiesysteem	n.s.	
Gebruik van testen	n.s.	
Onderwijskundig leiderschap	n.s.	
Nadruk cognitieve doelen	n.s.	
Teamcohesie	Groep B, D >	
Stimuleringskenmerken school		
Ouder participatie	Groep D >	
Intercultureel onderwijs	Groep D, E >	
OETC	Groep D, E >	
Instructie effectiviteit		
Klassikaal werken	Groep D >	*
Ordelijk klasklimaat	n.s.	
Minimumdoelen taal	Groep D >	*
Minimumdoelen rekenen	Groep D >	*
Feedback	n.s.	
Herhaling leerstof	n.s.	
Registratie vorderingen	n.s.	
Gebruik van testen	n.s.	
Huiswerk	Groep D >	*
Nadruk op basisvaardigheden	Groep D >	*
Stimuleringskenmerken klas		
Ouderparticipatie	Groep D >	*
Extra materiaal	Groep D >	
Projectonderwijs	n.s.	
Intercultureel onderwijs	Groep D >	*
Frequentie ICO	Groep D >	*
Communicatief taal	Groep D >	
Contacten met ouders	Groep D <	
Gezinsactivering	Groep D <	
Taalcompensatie	Groep D >	

De prestaties van de OVB-doelgroepleerlingen in 1990: stand van zaken na de tweede meting

P. Tesser en H. Vierke

Samenvatting

Ten behoeve van de evaluatie van het Onderwijsvoorrangsbeleid voeren het ITS en het RION een reeks longitudinale onderzoeken uit naar de ontwikkeling van de schoolprestaties van opeenvolgende cohorten basisschoolleerlingen. De eerste meting heeft in 1988 plaatsgevonden, de tweede meting in 1990. Dit artikel gaat over de resultaten van beide metingen. Het begint met enkele analyses van de gegevens uit de tweede meting. Vervolgens beschrijft het cross-sectionele en longitudinale vergelijkingen van de resultaten van 1988 en 1990. De achterstand van de leerlingen uit de doelgroepen van het onderwijsvoorrangsbeleid blijkt zowel cross-sectioneel als longitudinaal in 1990 precies even groot als in 1988. Dit geldt in gelijke mate voor scholen die in het geheel niet, in beperkte zin of intensief bij het OVB zijn betrokken.

1 Inleiding

Het onderwijsvoorrangsbeleid (OVB) is gericht op groepen leerlingen bij wie als gevolg van hun sociaal-etnische achtergrond de schoolprestaties en de schoolloopbanen achterblijven. Met het OVB wil de overheid de schoolprestaties en de schoolloopbanen van deze leerlingen verbeteren. De vraag is of en in hoeverre dit lukt en welke factoren hierbij een rol spelen.

Om hier zicht op te krijgen zijn het ITS en het RION in 1987 gestart met de uitvoering van een breed opgezet onderzoeksprogramma, de landelijke evaluatie van het onderwijsvoorrangsbeleid (LEO-OVB) (Kloprogge, 1986). De kern van het programma bestaat uit een serie longitudinale onderzoeken bij opeenvolgende leerlingcohorten in groep 4, 6 en 8 van het basisonderwijs en bij de scholen waar deze leerlingen op zitten. De longitudinale onderzoeken verlopen in fases: om de twee jaar worden gegevens verzameld over scholen, leerlingen en de ouders. De longitudinale onderzoeken hebben tot doel de ontwikkeling van de schoolprestaties en algemene cognitieve vaardigheden van de cohortleerlingen vast te stellen. Daartoe maken de leerlingen een taalrekentoets en een intelligentietest. De schooldirecteuren en leerkrachten beantwoorden schriftelijk vragen over het functioneren van de school en over hun werkwijze in de klas. De ouders van de leerlingen krijgen een kleine lijst met vragen over schoolondersteunend gedrag en de beschikbaarheid van culturele hulpbronnen in het gezin.

De longitudinale onderzoeken geven in de eerste plaats informatie over de schoolprestaties en de verdere loopbanen van de leerlingen uit de doelgroepen van het onderwijsvoorrangsbeleid. De belangrijkste doelgroepen zijn autochtone en allochtone leerlingen van wie tenminste een van beide ouders een laag opleidings- of beroepsniveau heeft. De leerlingen uit deze doelgroepen tellen bij de bepaling van de formatie van basisscholen zwaarder mee in het leerlingenaantal van de school dan leerlingen die niet tot de doelgroepen horen. De extra formatie die scholen voor de doelgroepleerlingen krijgen, is een van de twee hoofdcomponenten van het OVB. De tweede component bestaat uit de voorrangsgebieden. Dit zijn plaatselijke of regionale samenwerkingsverbanden tussen scholen en welzijnsinstellingen die tot doel hebben concrete vormen van aanpak van onderwijsachterstand tot ontwikkeling te brengen. De gebieden beschikken over middelen voor coördinatie en voor specifieke door scholen uit te voeren projecten.

Als het OVB effect heeft dan zullen de prestaties van de leerlingen uit de doelgroepen op den duur dichter in de buurt komen van de gemiddelde prestaties van alle basisschoolleerlingen. Het uiteindelijke doel is de achterstand van de doelgroepleerlingen op te heffen dan

wel te verminderen. In de longitudinale onderzoeken proberen we vast te stellen of een dergelijke ontwikkeling zich inderdaad voordoet. Het gaat hierbij in de eerste plaats om het opsporen van cohorteffecten. Doen de doelgroepleerlingen die in 1990 in groep 4, 6 of 8 zitten het beter dan de doelgroepleerlingen die in 1988 in diezelfde jaargroepen zaten? Daarnaast zijn we geïnteresseerd in longitudinale effecten: hoe zijn de prestaties van de doelgroepleerlingen die in 1988 in groep 4 zaten twee jaar later in groep 6?

De eerste fase van de longitudinale onderzoeken is in het schooljaar 1988/1989 uitgevoerd. De resultaten gaven een beeld van de taal- en rekenachterstand van de doelgroepleerlingen in 1988. Omdat het een beginmeting was, konden nog geen uitspraken worden gedaan over eventuele effecten van het beleid. Van deze eerste fase wordt wat het leerlinggedeelte betreft, verslag gedaan in de bijdrage van Van der Werf, Mulder en Tesser in dit nummer.

De belangrijkste resultaten van de eerste fase waren dat de prestaties van de allochtone doelgroepleerlingen veel slechter zijn dan we op grond van vooronderzoek (Tesser, Van der Werf, Mulder & Weide, 1989) hadden verwacht en dat de criteria voor de autochtone doelgroep niet stringent genoeg zijn. Veel van deze leerlingen zijn niet te beschouwen als kinderen met een maatschappelijke achterstand. Verder bleek dat de prestaties op scholen in voorrangsgebieden slechter zijn en dat dit een gevolg is van het hoge percentage allochtone leerlingen op deze scholen.

Inmiddels is ook de tweede fase van de longitudinale onderzoeken afgerond. Het is nu mogelijk de prestaties van de doelgroepleerlingen in de eerste fase te vergelijken met de prestaties in de tweede fase. Dit geeft zicht op mogelijke effecten van het OVB.

In dit artikel presenteren we gegevens ter beantwoording van de volgende twee algemene vragen:

1 Hoe staat het in 1990 met de prestaties in taal en rekenen van de leerlingen uit de belangrijkste OVB-doelgroepen en in hoeverre zijn er hierin verschillen met de situatie in 1988?
2 In hoeverre zijn er in de vergelijking van de resultaten van 1988 en 1990 effecten van OVB te onderkennen?

Om de beantwoording van de tweede vraag handen en voeten te geven hebben we scholen ingedeeld naar de mate van hun betrokkenheid bij het OVB.

2 Steekproeven en variabelen

Bij de eerste toetsafname in schooljaar 1988/1989 waren bijna 40.000 leerlingen betrokken. De leerlingen maakten deel uit van drie cohorten: 88-4, 88-6 en 88-8. De cohorten bestaan uit leerlingen die in het schooljaar 1988/1989 in groep 4, 6 of 8 zaten.

De tweede fase van de longitudinale onderzoeken is in het schooljaar 1990/1991 uitgevoerd. In dat schooljaar zijn de leerlingen uit de cohorten 88-4 en 88-6 die in groep 6 en 8 werden aangetroffen opnieuw onderzocht, tegelijk met de andere leerlingen in die jaargroepen. Uit de 'oude' 88-cohorten zijn de leerlingen verdwenen die zijn blijven zitten of verhuisd of die om andere redenen niet zijn doorgestroomd naar de hogere jaargroepen van dezelfde school. Verder is in 1990/1991 een nieuw groep 4-cohort onderzocht, cohort 90-4.

In de tweede fase zijn dus twee (deel)cohorten uit de eerste fase te onderscheiden: 88-4 en 88-6, en drie complete cohorten uit 1990: 90-4, 90-6 en 90-8. Alleen over de eerste twee cohorten, 88-4 en 88-6 zijn er longitudinale gegevens. Elk cohort bestaat uit ongeveer 13.000 leerlingen. De 'longitudinale cohorten' 88-4 en 88-6 bestaan elk uit ongeveer 8.000 leerlingen.

In de OVB-evaluatie werken we in hoofdzaak met drie categorieën leerlingen. De leerlingcategorieën zijn gebaseerd op de wegingsfactoren van de OVB-formatieregeling: allochtone doelgroepleerlingen, autochtone doelgroepleerlingen en leerlingen die niet tot de OVB-doelgroepen horen. Kortheidshalve duiden we de verschillende categorieën aan met de wegingsfactor: 1.90, 1.25 en 1.00. De doelgroepen met de wegingsfactor 1.40 en 1.70, de schipperskinderen en de woonwagenkinderen blijven hier verder buiten beschouwing. Leerlingen die tot deze groepen behoren hebben we ingedeeld bij de groep met de wegingsfactor 1.25. Het gaat om enkele tientallen leerlingen.

De cohortsteekproeven van de landelijke evaluatie zijn gedeeltelijk getrapte steekproeven. Eerst is een steekproef van scholen getrokken en vervolgens zijn binnen deze scholen alle leerlingen die voldoen aan de cohortdefinitie bij het onderzoek betrokken. In totaal nemen ruim 600 scholen deel aan de longitudinale onderzoeken. De scholensteekproef bestaat uit twee hoofdbestanddelen: een referentiesteekproef en een aanvullende steekproef. De referentiesteekproef is een toevalsselectie van 3.5 procent van alle Nederlandse basisscholen. Deze scholen zijn representatief voor de basisscholen in Nederland. De referentiesteekproef fungeert in de cohortonderzoeken als normgroep waartegen de prestaties van leerlingen uit de OVB-doelgroepen worden afgezet.

In de referentiesteekproef zitten te weinig scholen die voor de OVB-evaluatie van belang zijn. De referentiesteekproef bevat namelijk slechts enkele scholen met veel 1.90- en/of 1.25-leerlingen en ook zitten er weinig scholen in die deelnemen aan een onderwijsvoorrangsgebied. Daarom is deze steekproef aangevuld met alle scholen uit de helft van de onderwijsvoorrangsgebieden en met scholen die niet deelnemen aan een voorrangsgebied maar die toch veel 1.90 of 1.25-leerlingen hebben.

De scholen hebben we naar de aard van hun betrokkenheid bij OVB in drie typen verdeeld: gebiedsscholen, formatiescholen en niet OVB-scholen. Gebiedsscholen krijgen OVB-formatie en participeren aan een voorrangsgebied. De meeste gebiedsscholen hebben relatief veel 1.90-leerlingen. Formatiescholen krijgen wel extra formatie maar maken geen deel uit van een voorrangsgebied. Deze scholen hebben gemiddeld meer 1.25-leerlingen dan de gebiedsscholen en de scholen zonder OVB-formatie. Scholen zonder OVB-formatie hebben geen of vrijwel geen 1.90-leerlingen en merken in feite weinig van het OVB; ze kunnen nog wel een behoorlijk percentage 1.25-leerlingen hebben.

Van alle Nederlandse basisscholen is ruim veertig procent een formatieschool, zo'n negen procent is een gebiedsschool, iets meer dan 50 procent van de scholen valt buiten het OVB.

3 De vergelijkbaarheid van de toetsscores

Voor het vaststellen van de ontwikkeling van de taal- en rekenprestaties van de leerlingen uit de OVB-doelgroepen zijn bij de leerlingen speciaal voor de longitudinale onderzoeken ontworpen toetsen afgenomen. De toetsen zijn ontwikkeld door het CITO (Van Bergen, 1989). Als scores op de toetsen gelden de aantallen goed gemaakte opgaven op de toetsen. Dit levert echter enkele problemen op. Er zijn voor de onderscheiden jaargroepen verschillende toetsen, de ruwe scores op deze toetsen zijn niet met elkaar vergelijkbaar. De toetsen verschillen tussen de jaargroepen en naar vakgebied, de aantallen opgaven zijn ongelijk en de aantallen antwoordalternatieven variëren. Hier komt nog bij dat de moeilijkheidsgraad niet steeds hetzelfde is. Het gemiddelde aantal goed gemaakte opgaven van bijvoorbeeld de 1.25-leerlingen in groep 4 is daardoor niet zonder meer vergelijkbaar met het gemiddelde aantal van de 1.25-leerlingen in groep 6. Vanwege de verschillen in de aantallen antwoordalternatieven en de verschillen in moeilijkheidsgraad is ook het percentage goed gemaakte opgaven geen geschikte maat voor de prestaties. De onvergelijkbaarheid van de toetsscores in de verschillende jaargroepen is in ons geval extra lastig omdat we ook longitudinale vergelijkingen van de resultaten van verschillende groepen leerlingen willen uitvoeren. We willen bijvoorbeeld weten of de achterstand van de 1.90-leerlingen die in 88 in groep 4 zaten twee jaar later in groep 6 groter of kleiner is geworden.

Om dit probleem van de onvergelijkbaarheid van de toetsresultaten zo goed mogelijk te ondervangen hebben we de aantallen goed gemaakte opgaven van de leerlingen op de verschillende toetsen omgezet in relatieve scores. De scores van de individuele leerlingen zijn in oplopende volgorde gerangschikt. De zo verkregen rangen zijn getransformeerd tot normaalscores, met een gemiddelde van 0 en een standaarddeviatie van 1. Van deze normaalscores zijn vervolgens T-scores gemaakt met een gemiddelde van 50 en een standaarddeviatie van 10: $T = 10 * Z + 50$.
Hierin is:
T = T-score met een gemiddelde van 50 en een

standaarddeviatie van 10;
Z = de genormaliseerde rangscores.

We zijn in de longitudinale onderzoeken in de eerste plaats geïnteresseerd in de achterstand die de leerlingen uit de OVB-doelgroepen hebben ten opzichte van het landelijk gemiddelde. Daarom hebben we de normaalscores in eerste instantie berekend voor de leerlingen in de referentiesteekproef. Vervolgens is aan de somscore van de overige leerlingen de T-score toegekend die overeenkomt met die voor deze somscore berekende T-score in de referentiesteekproef.

Door deze operatie hebben de scores in de referentiesteekproef op elke toets in elke jaargroep een gemiddelde van 50 en een standaarddeviatie van 10. Iets meer dan 15 procent van de leerlingen in de referentiesteekproef heeft een score kleiner dan 40. Een even groot percentage heeft een score groter dan 60. Iets minder dan 35 procent heeft een score tussen 40 en 50 en eveneens iets minder dan 35 procent zit tussen 50 en 60.

Om enig idee te geven wat de betekenis is van de T-scores in relatie tot de aantallen goed gemaakte opgaven staan in Tabel 1 enkele kerncijfers. De tabel bevat de aantallen goedgemaakte opgaven op de taal- en rekentoets in 1990 in de groepen 4, 6 en 8 die corresponderen met T-scores van 40, 50 en 60. In de tabel hebben we ook nog een keer het totaal aantal toetsopgaven opgenomen.

Tabel 1

Correspondentie tussen T-scores en aantallen opgaven goed op de taal- en rekentoets in de groepen 4, 6 en 8 in 1990. Aantallen goed bij T-scores van 40, 50 en 60

	aantallen opgaven goed bij			maximum
	T = 40	T = 50	T = 60	goed
TAAL				
groep 4	44	53	59	65
groep 6	38	45	53	60
groep 8	48	55	62	67
REKENEN				
groep 4	27	33	38	40
groep 6	22	29	36	40
groep 8	21	28	35	40

De T-scores zeggen iets over de *relatieve* prestaties van afzonderlijke leerlingen of groepen leerlingen op taal en rekenen in een bepaalde jaargroep. Met de T-scores kunnen we deze prestaties vergelijken met die van een soortgelijke groep leerlingen in een andere jaargroep of met de prestaties van diezelfde leerlingen twee jaar later.

Uitspraken over verschillen tussen cohortpopulaties als geheel in termen van leerwinst zijn met de T-scores niet mogelijk omdat het gemiddelde in elke populatie steeds hetzelfde is. Omdat het in de eerste plaats gaat om de beschrijving en vergelijking van de prestaties van specifieke groepen leerlingen, de doelgroepen van het OVB, is dit niet zo'n groot bezwaar.

4 De taal- en rekenprestaties in 1990

Om te beginnen laten we de taal- en rekenprestaties in 1990 zien van de 1.90-, de 1.25- en de 1.00-leerlingen in de groepen 4, 6 en 8 van het basisonderwijs. We beschrijven de resultaten van de verschillende categorieën aan de hand van de gemiddelden en de standaarddeviaties van de T-scores. De cijfers staan in Tabel 2.

Tabel 2

Taal- en rekenprestaties in 1990 van de 1.90-, de 1.25 en de 1.00-leerlingen in groep 4, 6 en 8. Gemiddelden (gem) en standaardafwijkingen (sd) van de T-scores

	1.90		1.25		1.00	
	gem	*sd*	*gem*	*sd*	*gem*	*sd*
TAAL						
groep 4	41.3	9.4	48.5	9.6	52.2	9.6
groep 6	42.0	8.7	48.4	9.4	52.4	9.5
groep 8	40.8	9.2	48.5	9.6	52.8	9.3
REKENEN						
groep 4	44.3	9.7	48.6	9.9	51.6	9.6
groep 6	44.4	9.7	48.5	9.7	51.8	9.7
groep 8	43.3	9.6	48.2	9.7	52.6	9.4

Deze tabel is een toonbeeld van regelmaat. In alle drie de jaargroepen liggen de taal- en rekenscores van de leerlingen uit de OVB-doelgroepen, de 1.90- en de 1.25-leerlingen, onder het landelijk referentiegemiddelde. De scores van de 1.00-leerlingen liggen in alle gevallen boven het landelijk gemiddelde.

De scores van de 1.25-leerlingen zitten weliswaar onder het landelijk gemiddelde, maar de verschillen ten opzichte hiervan zijn gering: in

de meeste gevallen zo'n anderhalf punt. De 1.90-leerlingen komen veel lager uit. Op taal ligt hun gemiddelde score 8 tot 10 punten onder het landelijk referentiegemiddelde; op rekenen 5 tot 6 punten.

Bij vergelijking van de drie leerlinggroepen onderling zien we grote verschillen. Op taal scoren de 1.90-leerlingen 6 tot 8 punten lager dan de 1.25-leerlingen. De 1.25-leerlingen blijven op hun beurt op taal gemiddeld ongeveer 4 punten achter bij de 1.00-leerlingen.

Op rekenen is het verschil tussen de 1.90-leerlingen en de 1.25-leerlingen ongeveer 4 punten. Het verschil tussen de 1.25-leerlingen en de 1.00-leerlingen is zo'n 3 punten.

Bij de 1.00 en de 1.25-leerlingen is weinig verschil tussen de scores op taal en rekenen. Bij de 1.90-leerlingen ligt dit anders. Hun gemiddelde rekenscore ligt beduidend hoger dan hun gemiddelde taalscore. Dit wijst erop dat bij de 1.90-leerlingen vooral taal het grote knelpunt is. Rekenen leer je meer dan taal in hoofdzaak op school. Je zou kunnen zeggen dat de rekenscore aangeeft hoeveel minder de allochtone leerlingen van de schoolse component van het leren oppikken. In de taalscore komt daar nog een buitenschoolse component bovenop. Hierbij moeten we wel bedenken dat er ook in het rekenonderwijs, zeker in de hogere jaargroepen een taalcomponent zit.

5 Verschillen tussen groep 4, 6 en 8

Bij nauwkeurige bestudering van de cijfers in Tabel 2 tekent zich een tendens af naar grotere verschillen tussen de leerlingcategorieën in de hogere leerjaren. Met name de achterstand van de 1.90-leerlingen lijkt in de hogere jaargroepen iets groter. Om hier duidelijkheid over te krijgen hebben we de verschillen tussen leerlingcategorieën en jaargroepen verder onderzocht met behulp van variantie-analyse. Hiertoe zijn de gegevens van de drie jaargroepen bij elkaar gevoegd.

In de variantie-analyse die we hebben uitgevoerd zijn wegingsfactor en jaargroep factoren die mogelijk een deel van de verschillen in toetsprestaties kunnen verklaren. In de analyse toetsen we of de factoren ieder afzonderlijk of combinaties van hun categorieën, hun interactie, een significante eigen bijdrage leveren aan de verklaring van verschillen tussen de leerlingen op de taal- en rekentoetsen.

Als de tendens in Tabel 2, namelijk dat de verschillen tussen de leerlingcategorieën in de hogere leerjaren groter zijn, zich inderdaad voordoet dan moet dit in de variantie-analyse tot uiting komen in een significant interactie-effect van jaargroep met wegingsfactor. Dat wil zeggen dat bepaalde combinaties van wegingsfactor en leerjaar een andere score opleveren dan je zou krijgen door de effecten van die twee factoren gewoon op te tellen.

Toetsing van de effecten gebeurt in de variantie-analyse aan de hand van de grootheid F. F is de verhouding tussen de verschillen die gebonden zijn aan een bepaalde variantiecomponent enerzijds en de niet gebonden variantie anderzijds. De kansverdeling van F is bekend en als de kans dat een gevonden F aan toeval te wijten is, erg klein is dan is er sprake van een significant, dat wil zeggen betekenisvol effect.

De resultaten van de variantie-analyse staan in Tabel 3. Deze tabel bevat de verschillen, of preciezer, de sommen van de gekwadrateerde verschillen die gebonden zijn aan de factoren: de kwadratensommen. Verder staan in Tabel 3 de aantallen vrijheidsgraden per analysecomponent, de waarde van F voor die component en de kans om die F bij toeval te vinden. Het aantal vrijheidsgraden is het aantal categorieën waarover de verschillen die gebonden zijn aan een bepaalde component vrij zijn te verdelen. Het aantal vrijheidsgraden speelt een rol bij de berekening van F en bij de bepaling van de kanswaarde voor F.

Uit de analyses blijkt dat er zowel bij taal als bij rekenen significante interactie-effecten zijn van wegingsfactor en jaargroep. In dit geval betekent dit dat de verschillen tussen leerlingen met een verschillende wegingsfactor in de hogere leerjaren van het basisonderwijs in 1990 significant groter zijn dan in de lagere leerjaren.

Hoewel de verschillen significant zijn, kunnen we daar toch weinig waarde aan hechten. Het interactie-effect verklaart zowel bij taal als ook bij rekenen slechts eentiende procent van de totale variantie, en is daarom te verwaarlozen.

Tabel 3

Variantie-analyse van taal-(T) en rekenscores (R) in 1990, met wegingsfactor en jaargroep als hoofdeffecten: kwadratensommen, vrijheidsgraden (df), F-waarden en kansen op F (p)

variantiebronnen		kwadratensom	*df*	*F*	*p*
wegingsfactor	(T)	583579.9	2	3349.5	.00
	(R)	300004.9	2	1600.2	.00
jaargroep	(T)	312.2	2	1.8	.17
	(R)	783.6	2	4.2	.02
interactie	(T)	2408.5	4	6.9	.00
	(R)	3322.5	4	8.9	.00
niet gebonden	(T)	2752230.4	31593		
	(R)	2961495.4			

Tabel 4

Gemiddelde t-scores op taal (T) en rekenen (R) per schooltype in 1990: groep 4, 6 en 8

Schooltype	groep 4		groep 6		groep 8	
	(T)	(R)	(T)	(R)	(T)	(R)
gebied	43.8	45.6	43.6	45.4	43.3	44.8
formatie	49.1	48.6	48.7	48.7	48.6	48.5
niet-OVB	51.1	51.1	51.5	51.2	51.9	51.6

Betekent dit nu ook dat de verschillen groter worden naarmate de leerlingen verder in het onderwijs komen? Het ziet er wel naar uit, maar op grond van deze analyse alleen kunnen we dat strikt genomen nog niet zeggen. De leerlingen in groep 4 zijn andere dan in groep 6 en 8 en het kan dus ook zijn dat de doelgroepleerlingen in de jongere jaargroepen het beter doen, bijvoorbeeld omdat de aanpak van achterstand in de eerste plaats is gericht op de lagere jaargroepen. Ook is het mogelijk dat de jongere cohorten om andere redenen beter zijn dan de oudere. Om zicht te krijgen op de ontwikkeling van de prestaties van de verschillende groepen leerlingen hebben we longitudinale analyses nodig. Die behandelen we verderop in dit artikel.

6 Verschillen tussen schooltypen

Het OVB werkt via scholen en voorrangsgebieden. Scholen krijgen extra formatie en gebieden proberen inhoud te geven aan voorrang door de ontwikkeling en toepassing van speciale programma's, projecten en activiteiten voor leerlingen uit de doelgroepen. In de cohortonderzoeken proberen we zicht te krijgen op de effecten van een en ander door prestatieverschillen te analyseren tussen relevante categorieën leerlingen en tussen schooltypen: gebiedsscholen, formatiescholen en scholen zonder OVB-formatie.

Hierboven hebben we gezien dat ook in 1990 de leerlingen uit de OVB-doelgroepen over het geheel genomen nog duidelijk lagere taal- en rekenprestaties halen dan de leerlingen die niet tot de doelgroepen horen. Maar hoe staat het met de schooltypen?

Tabel 4 geeft de gemiddelde toetsscores per schooltype. De standaarddeviaties zijn ter wille van de overzichtelijkheid weggelaten. Ze variëren tussen 9.6 en 10.2 en wijken dus in elke afzonderlijke groep minimaal af van de standaardafwijking in de totale jaargroepen.

De niet-OVB-scholen halen in alle drie de jaargroepen zowel op taal als op rekenen duidelijk de hoogste gemiddelden. De formatiescholen zitten daar zo'n 2 tot 3 punten onder. De gemiddelden van de gebiedsscholen zijn veel lager. Ze liggen ongeveer 6 tot 8 punten onder de gemiddelden van de niet-OVB-scholen.

De verhoudingen tussen de schooltypen lijken enigszins op de verhoudingen tussen de leerlingcategorieën (Tabel 2). De scholen zonder OVB-formatie zitten ongeveer op het niveau van de 1.00-leerlingen, de formatiescholen hebben de positie van de 1.25-leerlingen en de gebiedsscholen komen op de plaats van de 1.90-leerlingen. Dit is niet zo vreemd omdat deze leerlingcategorieën op de betreffende schooltypen sterk zijn oververtegenwoordigd. Zo bestaat de leerlingbevolking van OVB-gebiedsscholen voor bijna 40 procent uit 1.90-leerlingen, terwijl dit op niet-OVB-scholen nog geen 2 procent is.

Maar betekent dit nu dat de verschillen tussen de schooltypen geheel zijn te herleiden tot verschillen in de samenstelling van de schoolbevolking? Of dit het geval is, hebben we weer uitgezocht met variantie-analyse. Als de verschillen tussen schooltypen volledig het gevolg zijn van verschillen in samenstelling van de leerlingbevolking dan moeten in variantie-analyses met wegingsfactor en schooltype als onafhankelijke verklarende factoren alleen significante hoofdeffecten van de wegingsfactor te vinden zijn en geen significante effecten van schooltype. De betreffende analyses zijn uitgevoerd over de gegevens van groep 4, 6 en 8. De resultaten zijn in elke jaargroep ongeveer hetzelfde, daarom beperken we de presentatie ervan tot groep 8. De analysegegevens staan in Tabel 5. Deze tabel begint met de gemiddelde toetsscores per combinatie van wegingsfactor en schooltype. Daarna volgen de resultaten van de variantie-analyse met de taal- en rekenscore als afhankelijke variabelen en wegingsfactor en schooltype als verklarende factoren.

Aan de gemiddelden is al te zien dat ook bij gelijke wegingsfactor de leerlingen op de niet-OVB-scholen het beter doen dan de leerlingen op de OVB-scholen. In alle gevallen zijn de scores van de leerlingen op de niet-OVB-scholen het hoogst. De indruk die uit de gemiddelden naar voren komt wordt bevestigd door de variantie-analyse. De hoofdeffecten, wegingsfactor èn schooltype, zijn beide significant. Alle leerlingen halen ongeacht hun wegingsfactor op OVB-gebiedsscholen zwakkere resultaten dan op OVB-formatiescholen. Op de formatiescholen zijn de resultaten weer minder dan op de niet-OVB-scholen.

De gemiddelden van leerlingen met een bepaalde wegingsfactor op een bepaald schooltype zijn niet volledig de optelsom van de effecten van schooltype en wegingsfactor. Dit komt doordat er ook significante interactie-effecten zijn. Deze zijn echter erg klein en er zit geen duidelijke lijn in. Daarom laten we ze hier verder buiten beschouwing.

Leerlingen op OVB-scholen halen slechtere resultaten dan leerlingen op niet-OVB-scholen. Moeten we hieruit afleiden dat het OVB niet of zelfs averechts werkt? De verschillen tussen schooltypen waren er ook bij de eerste meting in 1988. Bij nadere analyse bleek toen dat de verschillen niets te maken hadden met de deelname van de scholen aan het beleid. De prestaties op de OVB-scholen waren lager omdat de sociale achtergrond van de leerlingen op deze scholen ook bij gelijke wegingsfactor ongunstiger was dan bij leerlingen op niet-OVB-scholen èn omdat er op de OVB-scholen veel meer 1.90-leerlingen zitten. In Tesser en Mulder (1990) hebben we laten zien dat het percentage 1.90-leerlingen op een school een negatief effect heeft op de gemiddelde prestaties.

Omdat de oorzaken van de verschillen tussen de schooltypen van groot belang zijn voor een beoordeling van de effecten van het beleid hebben we ook onderzocht hoe dit in 1990 zit.

Tabel 5

Variantie-analyses van de taal- (T) en rekenscores (R) in 1990 in groep 8 met wegingsfactor en schooltype als hoofdeffecten: gemiddelden per combinatie van wegingsfactor en schooltype, kwadratensommen, vrijheidsgraden(df), F-waarden(F) en kansen op F(p)

		1.90		1.25		1.00	
		(T)	(R)	(T)	(R)	(T)	(R)
gebied		39.5	42.5	45.8	45.9	51.1	50.7
formatie		41.8	43.4	47.5	47.3	51.9	51.6
niet-OVB		42.9	46.2	50.0	49.4	53.2	53.0
variantiebronnen	(T) (R)	kwadratensom		*df*		*F*	*p*
wegingsfactor	(T)	328338.8		2		1878.1	.00
	(R)	238445.0		2		1326.2	.00
schooltype	(T)	52465.5		2		300.1	.00
	(R)	44185.9		2		245.7	.00
interactie	(T)	5202.8		4		14.9	.00
	(R)	1917.7		4		5.3	.00
niet gebonden	(T)	4223129.3		48313			
	(R)	4343388.8					

7 Sociale achtergrond, samenstelling leerlingbevolking en prestaties

In hoeverre zijn in 1990 de verschillen tussen de schooltypen toe te schrijven aan de sociale achtergrond van de leerlingen en de etnische samenstelling van de school? Om hier een beeld van te krijgen hebben we de variantie-analyses met wegingsfactor en schooltype van Tabel 5 nog een keer uitgevoerd, nu met controlevariabelen voor sociale achtergrond en samenstelling van de schoolbevolking. Als indicatoren voor de sociale achtergrond hebben we de opleiding van de vader en de moeder van de leerlingen genomen. Hierbij moeten we wel een kanttekening plaatsen. Gegevens omtrent de sociale achtergrond ontbreken bij in totaal 29 procent van de leerlingen. Uitgesplitst naar wegingsfactor ontbreken deze gegevens bij 28 procent van de 1.00-leerlingen en bij 26 procent van de 1.25-leerlingen. Bij de 1.90-leerlingen is de uitval hoger: 44 procent. De etnische samenstelling van de schoolbevolking is gemeten aan de hand van het percentage 1.90-leerlingen in de groep. Opleidingsniveau van de ouders en percentage 1.90-leerlingen zijn in de variantie-analyse als covariaten opgenomen.

In de variantie-analyse met covariaten worden de taal- en rekenscores per wegingsfactor en schooltype gecorrigeerd voor de effecten van opleiding vader en moeder en van het percentage allochtone leerlingen in de groep. We hebben eerst gecontroleerd voor de wegingsfactor en vervolgens voor de covariaten, opleidingsniveau van de ouders en percentage 1.90-leerlingen in de klas. De analyse levert verschillen in gemiddelden op tussen de schooltypen, eerst bij gelijke wegingsfactor en vervolgens bij gelijke wegingsfactor, gelijk opleidingsniveau van de ouders en gelijk percentage 1.90-leerlingen. De mate waarin de correctie leidt tot een vermindering van de verschillen tussen schooltypen valt af te leiden uit een vergelijking van de analyseresultaten zonder en met de covariaten.

In Tabel 6 zijn alleen die resultaten van de analyse gepresenteerd die relevant zijn in verband met de vraag naar de verschillen tussen de schooltypen. In de tabel staan de gemiddelden van de schooltypen als afwijking van het algemeen gemiddelde van alle leerlingen. De tabel bevat ook het effect van schooltype in de vorm van een regressiecoëfficiënt (*bèta*). Uit de verschillen tussen de regressiecoëfficiënten valt af te leiden welk deel van de verschillen tussen de schooltypen is toe te schrijven aan de covariaten.

Wanneer we bij het bekijken van de cijfers in deze tabel ook nog de gegevens uit Tabel 5 betrekken, die de 'ruwe' verschillen tussen de schooltypen geeft, kan precies worden getraceerd hoe deze verschillen zijn opgebouwd. Zo is op taal het 'ruwe verschil' tussen gebiedsscholen en niet-OVB-scholen 8.6-punten. Bij gelijke wegingsfactor is dit nog 3.1 en na controle voor sociale achtergrond en percentage 1.90-leerlingen in de klas blijft er een verschil van 1.4-punten over (-.86 -.55 =-1.41).

In totaal komt ongeveer een kwart van de verschillen tussen de schooltypen voor rekening van de sociale achtergrond en nog eens een kwart voor rekening van het percentage 1.90-leerlingen in de groep.

De correctie met de covariaten leidt er vooral toe dat de verschillen tussen de gebiedsscholen en de formatiescholen vrijwel verdwijnen. Ook van de verschillen met de niet-OVB-scholen blijft weinig over. Het totale schooltype-effect is nog wel significant maar zo gering dat we er niet echt rekening mee hoeven te houden.

Tabel 6

Effecten van schooltype op de taal- en rekenscores in 1990 in groep 8 onder constanthouding van wegingsfactor en zonder (ZC) en met controle (MC) voor opleiding ouders en percentage 1.90-leerlingen in de groep

		gebied	formatie	niet-OVB	*bèta*
taal	ZC	-2.26	-.90	.82	.10
taal	MC	-.86	-.56	.55	.05
rekenen	ZC	-1.63	-.78	.67	.08
rekenen	MC	-.39	-.47	.33	.04

Al met al kunnen we vaststellen dat ook in 1990 leerlingen op de OVB-scholen slechter presteren dan leerlingen op niet-OVB-scholen maar dat dit niet mag worden uitgelegd als een gevolg van het beleid. De resultaten op de OVB-scholen zijn minder omdat de sociale achtergrond van de leerlingen op deze scholen ongunstiger is en omdat het percentage allochtone doelgroepleerlingen op OVB-scholen hoger is.

Ook bij adequate controle voor opleiding van de ouders zijn de gemiddelde prestaties van alle leerlingen lager naarmate het percentage 1.90-leerlingen hoger is. OVB heeft er ook in 1990 nog niet tot resultaat gehad dat dit effect is verdwenen.

8 Vergelijking van de resultaten uit 1988 met 1990

De OVB-evaluatie is longitudinaal opgezet. Om de twee jaar worden in de groepen 4, 6 en 8 toetsen afgenomen op dezelfde scholen. Dit biedt de mogelijkheid zowel cross-sectionele als ook longitudinale analyses uit te voeren.

Bij de cross-sectionele analyses worden de resultaten van leerlingen uit een bepaalde jaargroep vergeleken met die van leerlingen uit dezelfde jaargroep twee jaar eerder. Hierbij kunnen eventuele verschillen met betrekking tot de onderwijspositie van de doelgroepleerlingen tussen leerjaren worden onderzocht. Het nadeel van deze vergelijking is dat de samenstelling van de jaargroepen verschillend kan zijn.

Naast cross-sectionele gegevens hebben we ook de beschikking over longitudinale gegevens van twee cohorten: leerlingen die in 1988 in groep 4 c.q. in groep 6 zaten zijn in 1990 voor de tweede keer getoetst, zij zaten toen respectievelijk in groep 6 en groep 8. De longitudinale data maken het mogelijk de ontwikkeling van leerlingen tussen twee meetmomenten vast te stellen. Het nadeel is dat alleen leerlingen in de analyses zijn opgenomen die aan beide metingen hebben meegedaan, uitvallers (bijvoorbeeld zittenblijvers) blijven buiten beschouwing. Strikt genomen gaat het dus alleen om de ontwikkeling van de betere leerlingen.

8.1 1988-1990: cross-sectionele vergelijking

In de OVB-evaluatie proberen we op verschillende manieren vat te krijgen op de effecten van het beleid. Omdat we rekening willen houden met de mogelijkheid dat eventuele effecten zich pas op den duur zullen voordoen, zijn we vooral geïnteresseerd in de ontwikkeling van de schoolprestaties van de leerlingen uit de doelgroepen in de loop van de tijd. Nu, na de tweede fase, zijn de gegevens uit 1988 en 1990 bekend en kan dus een eerste vergelijking in de tijd worden gemaakt.

In deze paragraaf zetten we de resultaten van de leerlingcohorten 88-4, 88-6 en 88-8 naast die van de cohorten 90-4, 90-6 en 90-8. Van elk van de jaargroepcohorten hebben we de scores op dezelfde taal- en rekentoetsen. De scores van de 88-cohorten zijn genormaliseerd op de verdelingen in 1988, die van de 90-cohorten op de verdelingen in 1990. In beide jaren geven de scores dus de relatieve posities weer van de leerlingen ten opzichte van het landelijk gemiddelde in het betreffende jaar. De scores van de leerlingen in 1988 en 1990 zijn hiermee optimaal met elkaar vergelijkbaar en de gemiddelden per wegingsfactor geven aan hoe de positie van de leerlingen uit de doelgroepen van het OVB in 1988 zich verhoudt tot de positie van de doelgroepen in 1990. De betreffende gemiddelden voor taal en rekenen staan in Tabel 7.

Is er tussen 1988 en 1990 iets veranderd in de scores van de leerlingen uit de OVB-doelgroepen? Vergelijking van de paren van gemiddelden in 1988 en 1990 per jaargroep en wegingsfactor geeft hierover uitsluitsel. In vrijwel alle gevallen zijn de scores in 1990 hoger dan in 1988. De verschillen zijn echter gering. Het grootste verschil bedraagt één score-

Tabel 7

Gemiddelde T-scores voor taal en rekenen per leerlingcategorie in 1988 en 1990 in groep 4, 6 en 8

	1.90		1.25		1.00	
	1988	1990	1988	1990	1988	1990
TAAL						
groep 4	41.4	41.3	48.5	48.6	51.8	52.2
groep 6	41.7	42.0	48.0	48.4	52.4	52.4
groep 8	40.1	40.8	48.3	48.4	52.4	52.8
REKENEN						
groep 4	44.1	44.3	48.6	48.6	51.3	51.6
groep 6	43.4	44.4	48.1	48.5	52.2	51.8
groep 8	43.0	43.3	47.9	48.2	52.2	52.6

punt, het gemiddelde verschil is iets meer dan drietiende punt.

We hebben de verschillen tussen wegingsfactoren, jaargroepen en onderzoeksjaren getoetst met variantie-analyse. Het onderzoeksjaareffect is significant.

We zijn vooral geïnteresseerd in de interactie tussen onderzoeksjaar en wegingsfactor. Is deze interactie significant dan betekent dit dat er tussen de leerlinggroepen verschillen zijn in de mate waarin de scores in 1990 hoger zijn dan in 1988. Dat kan wijzen op een OVB-effect.

De analyse wijst uit dat het interactie-effect noch voor taal noch voor rekenen significant is. Het sterkste interactie-effect dat de analyse oplevert is dat van jaargroep met wegingsfactor. Dit effect was ook te zien bij de analyses van de gegevens over 1990 afzonderlijk. Nu blijkt dus dat het zich in beide onderzoeksjaren voordoet. Het interactie-effect van jaargroep en wegingsfactor zegt in dit geval dat in beide onderzoeksjaren de verschillen tussen de scores van de leerlingen met verschillende wegingsfactor groter zijn in de hogere jaargroepen. Opnieuw zien we dus dat de achterstand van de leerlingen uit de OVB-doelgroepen groter is naarmate de leerlingen verder in het onderwijs zijn gevorderd.

Het belangrijkste resultaat van de crosssectionele vergelijking is de vaststelling dat de achterstand van de leerlingen uit de OVB-doelgroepen in 1990 vrijwel even groot is als in 1988. Tussen de cohorten 88 en 90 is er nauwelijks sprake van enige ontwikkeling. Alleen de achterstand van de 1.90-leerlingen op rekenen is in 1990 een fractie geringer dan in 1988.

8.2 1988-1990: longitudinale vergelijking

Bij de longitudinale analyses van de gegevens gaat het om de vraag hoe de leerlingen die in 1988 zijn getest het er in 1990 afbrengen. Het betreft de leerlingen die in 1988 in groep 4 en 6 zaten. De longitudinale gegevens laten, nauwkeuriger dan bij de vergelijking tussen cohorten, een antwoord toe op de vraag hoe de verschillende leerlingcategorieën zich ontwikkelen naarmate de leerlingen verder in het onderwijs komen. Omdat het om dezelfde leerlingen gaat zijn verschillen tussen 1988 en 1990 te interpreteren als verschillen in ontwikkeling van de relatieve taal- en rekenprestaties. Voor een goed begrip van de resultaten is het van belang te benadrukken dat we het steeds hebben over de *relatieve* prestaties van de leerlingen ten opzichte van hun medeleerlingen in hetzelfde cohort. Verschillen geven dus geen absolute leerwinst aan.

De longitudinale gegevens maken tot op zekere hoogte ook betere uitspraken mogelijk over de verschillen tussen OVB-gebiedsscholen, formatiescholen en niet-OVB-scholen. De prestaties in 1988 kunnen dienen als controle voor vooraf bestaande verschillen tussen de schooltypen en de effecten van de schooltypen zijn zuiverder te bepalen door hier bij de analyse van de resultaten van 1990 rekening mee te houden.

Doordat we bij de taal- en rekentoetsen met genormaliseerde rangscores werken, zijn de gemiddelde scores die de leerlingen in 1988 hebben gehaald te vergelijken met hun scores in 1990. Deze gemiddelden staan, uitgesplitst naar jaargroep en wegingsfactor in Tabel 8.

Tabel 8

Gemiddelde toetsscores op taal en rekenen per wegingsfactor van de leerlingen die in 1988 in groep 4 en 6 en in 1990 in groep 6 en 8 aan het onderzoek hebben meegedaan

	1.90	1.25	1.00
TAAL			
groep 4(88)	41.6	48.8	51.6
groep 6(90)	42.0	48.3	52.3
groep 6(88)	41.3	48.1	52.4
groep 8(90)	41.5	48.3	52.7
REKENEN			
groep 4(88)	43.8	48.8	51.2
groep 6(90)	44.2	48.5	51.6
groep 6(88)	43.6	48.4	52.1
groep 8(90)	43.6	48.0	52.5

In deze tabel moeten we per wegingsfactor kijken naar de verschillen tussen de (verticale) paren van gemiddelden in 1988 en 1990.

De verschillen binnen de paren zijn klein. Het grootste verschil bedraagt zeventiende punt. Om te achterhalen of er enige systematiek in de verschillen zit, hebben we ze ook statistisch getoetst. Dit kan op verschillende manieren, maar die zijn niet allemaal even goed. Een op het eerste gezicht voor de hand liggende methode is variantie-analyse met de meting in 1990 als afhankelijke variabele en de meting in 1988 als covariaat. Dit leidt echter tot mislei-

dende resultaten (zie Mulder, Vierke & Tesser, i.v.).

De meest in aanmerking komende alternatieven voor de covariantie-analyse zijn een herhaalde metingen-analyse en een variantie-analyse van verschilscores. Omdat die twee methoden in het geval van twee metingen op hetzelfde neerkomen en interpretatie van de analyseresultaten bij de verschilscores wat eenvoudiger is hebben we voor deze laatste methode gekozen.

Per leerling is het verschil berekend tussen de score op taal en rekenen in 1990 en in 1988. Met deze verschilscores als afhankelijke variabelen zijn vervolgens variantie-analyses uitgevoerd. We zijn daarbij voor de keuze van de verklarende variabelen uitgegaan van de verschillende invalshoeken die we hebben gehanteerd bij het bekijken van de gegevens over 1990. Dit zijn: de OVB-wegingsfactor van de leerlingen en het schooltype waar ze op zitten. Er zijn vier analyses uitgevoerd, twee analyses over de leerlingen die in 1990 in groep 8 zaten en in 1988 in groep 6, en twee analyses over verschillen in taal- en rekenscores van de leerlingen die in 1990 in groep 6 en in 1988 in groep 4 zaten. De groepen leerlingen duiden we verder aan met de benaming van het cohort waar ze bij de eerste meting deel van uitmaakten: cohort 88-6 en 88-4.

Met de variantie-analyses van de verschilscores kunnen verschillende vragen worden beantwoord. In de eerste plaats of er verschillen zijn in de ontwikkeling van de relatieve taal- en rekenprestaties van leerlingen met een verschillende wegingsfactor. In de tweede plaats of er verschillen zijn in de prestatie-ontwikkeling op de verschillende schooltypen. Verder brengen de analyses effecten aan het licht van specifieke combinaties van wegingsfactor en schooltype. De gemiddelde verschilscores naar wegingsfactor en schooltype vormen de basisgegevens voor de variantie-analyses. Ze staan in Tabel 9.

Tabel 9

Gemiddelde verschilscores tussen 1990 en 1988 op taal en rekenen per wegingsfactor en schooltype in de cohorten 88-4 en 88-6

	cohort 88-4		cohort 88-6	
	(T)	(R)	(T)	(R)
WEGINGSFACTOR				
1.90	.65	.52	.34	.00
1.25	-.36	-.32	.24	-.32
1.00	.87	.41	.24	.39
SCHOOLTYPE				
gebied	.18	.16	-.17	.12
formatie	-.44	-.46	-.32	-.15
niet-OVB	.92	.52	.77	.20

Op basis van de gemiddelden in deze tabel kunnen nog geen conclusies worden getrokken over de effecten van wegingsfactor en schooltype op de verschilscores. Vooral bij wegingsfactor en schooltype treedt er een vermenging van effecten op. Op gebiedsscholen zitten bijvoorbeeld veel meer 1.90-leerlingen dan op de andere schooltypen. Dit werkt door in het gemiddelde van de gebiedsscholen. De variantie-analyses leveren onafhankelijke effecten op van wegingsfactor en schooltype.

Van de variantie-analyses vermelden we hier alleen de belangrijkste resultaten. De hoofdeffecten van wegingsfactor en schooltype zijn wisselend. Soms zijn ze net wel, soms niet significant. Deze effecten zijn uiterst klein en er zit geen duidelijke lijn in. De conclusie moet zijn dat er geen systematische verschillen zijn in de ontwikkeling van de relatieve prestaties op taal en rekenen tussen leerlingen met een verschillende wegingsfactor en tussen leerlingen op de onderscheiden schooltypen.

Van belang zijn vooral de interactie-effecten. Deze kunnen een eventuele differentiële ontwikkeling van OVB-doelgroepleerlingen op OVB-scholen aan het licht brengen. Welnu, niets in de resultaten wijst erop dat dit het geval is. De interactie-effecten zijn nergens significant. Het is dus niet zo dat de ontwikkeling van de prestaties van leerlingen met een bepaalde wegingsfactor op bepaalde schooltypen duidelijk afwijkt van die van andere leerlingen. Een 'OVB-effect' doet zich niet voor.

In de cross-sectionele analyses bleek, dat de resultaten van de leerlingen met een verschillende wegingsfactor verder uiteen gaan lopen naarmate ze hoger in de basisschool komen. In de longitudinale analyse zijn, tenminste over een periode van twee jaar, geen systematische verschillen te vinden in de ontwikkeling van de toetsscores in samenhang met de wegingsfactor. Omdat in de longitudinale analyse veel meer externe factoren onder controle worden gehouden, moeten we er voorlopig vanuit gaan dat zo'n differentiële ontwikkeling zich ook in-

derdaad niet voordoet. Voor het OVB betekent dit dat de achterstand die leerlingen uit de doelgroepen in 1988 in groep 4 en groep 6 hadden twee jaar later in groep 6 en 8 niet groter, maar ook niet kleiner is geworden. In feite is er in de periode van twee jaar tussen 1988 en 1990 in dit opzicht niets veranderd.

9 Conclusie

In het schooljaar 1986/87 is het onderwijsvoorrangsbeleid volledig gaan draaien. Het doel is de achterstand in taal- en rekenprestaties van de leerlingen uit de doelgroepen te verminderen. Uit de resultaten die we in dit artikel hebben gepresenteerd blijkt dat hier weinig schot in zit. Ook in het schooljaar 1990/1991 zijn er grote verschillen in de prestaties van de onderscheiden groepen leerlingen. De taal- en rekenscores van de allochtone leerlingen blijven ver achter bij die van de autochtone doelgroepleerlingen van het OVB. De autochtone doelgroepleerlingen halen op hun beurt weer beduidend lagere scores dan de leerlingen van wie de ouders een hoger opleidings- en beroepsniveau hebben.

De resultaten van 1990/1991 wijken nauwelijks af van die in 1988/1989. Er zijn geen aanwijzingen dat het OVB een ontwikkeling in de richting van minder achterstand voor de doelgroepen op gang heeft gebracht. De leerlingen uit de doelgroepen doen het in 1990 in de overeenkomstige jaargroepen niet beter dan in 1988. Ook is er geen vermindering in de achterstand van de doelgroepleerlingen die tussen 1988 en 1990 twee jaargroepen verder in het basisonderwijs zijn gekomen. Enigszins opmerkelijk is wel dat de achterstand ook niet groter is geworden.

Ten gevolge van het OVB krijgen scholen met veel leerlingen uit de doelgroepen meer personeelsformatie en scholen in voorrangsgebieden nemen deel aan activiteiten en projecten die er specifiek op zijn gericht achterstand te verminderen. Dit resulteert vooralsnog niet in betere prestaties van de leerlingen uit de doelgroepen op deze scholen. De ontwikkeling van de prestaties tussen 1988 en 1990 is op de OVB-formatie en -gebiedsscholen hetzelfde als op de scholen die geen OVB-faciliteiten krijgen. Van een 'OVB-effect' in de zin van gunstiger ontwikkeling van de prestaties van doelgroepleerlingen op OVB-gebiedsscholen of OVB-formatiescholen is geen sprake.

Eerdere analyses van de gegevens uit 1988 hebben uitgewezen dat de prestaties op scholen met veel allochtone leerlingen lager zijn dan op scholen met weinig of geen allochtone leerlingen. De resultaten die we in dit artikel hebben laten zien wijzen erop dat het in 1990 niet anders is. Scholen met veel allochtone leerlingen nemen vaak deel aan een onderwijsvoorrangsgebied. Dit heeft tot gevolg dat de resultaten op scholen waar het OVB de meeste impact heeft het slechtst zijn. De faciliteiten en activiteiten leiden er niet toe dat de extra 'handicap' die het hebben van veel leerlingen uit de allochtone doelgroepen met zich meebrengt wordt overwonnen.

Alles overziende zijn de resultaten die de vergelijking van de eerste en de tweede fase van de longitudinale onderzoeken oplevert van een welhaast ontmoedigende gelijkvormigheid. Het is al eerder geconstateerd, onderwijsachterstand verminderen is moeilijk (Tesser, 1992).

De hier gepresenteerde analyses werpen licht op een van de mogelijke oorzaken hiervan. Een van de redenen waarom het zo lastig is achterstand ten gevolge van sociaal-etnische kenmerken van leerlingen aan te pakken ligt waarschijnlijk in de door ons gesignaleerde grote onderlinge verschillen in prestaties binnen de doelgroepen van het OVB. In elk van de doelgroepen zijn er altijd substantiële aantallen leerlingen die het zo slecht nog niet doen en die zelfs beter presteren dan leerlingen die niet tot de doelgroepen worden gerekend. Voor de docent zit hier vrijwel zeker een belangrijke belemmering om te komen tot een doelgerichte aanpak. In strategieën voor aanpak van onderwijsachterstand is hier ons inziens tot nu toe onvoldoende rekening mee gehouden.

Literatuur

Bergen, J. B. A. M. van (1989). *Verantwoording constructie toetsen voor de evaluatie van het onderwijsvoorrangsbeleid.* Arnhem: CITO.

Kloprogge, J. (1986). *Uitvoeringsplan evaluatie onderwijsvoorrangsbeleid.* Den Haag: SVO.

Tesser, P., Werf, M. van der, Mulder, L., & Weide, M. (1989). *De beginfase van het onderwijsvoorrangsbeleid in het basisonderwijs. Verslag van het oriënterend onderzoek van de landelijke OVB evaluatie.* Nijmegen/Groningen: ITS/RION.

Tesser, P., & Mulder, L. (1990). Concentratie en prestatie van allochtone leerlingen in het basisonderwijs. *Migrantenstudies, 2,* 31-44.

Tesser, P. (1992). *OVB-activiteiten ingekaderd. Aanbevelingen en checklist.* Nijmegen: ITS.

Manuscript aanvaard 11-1-1993

Auteurs

P. Tesser (1946) werkte tot 1992 op het Instituut voor Toegepaste Sociale Wetenschappen in Nijmegen en was daar onder andere projectleider van de Landelijke Evaluatie Onderwijsvoorrangsbeleid. Is sinds juni 1992 werkzaam bij het Sociaal Cultureel Planbureau.

H. Vierke (1953) is werkzaam bij het Instituut voor Toegepaste Sociale Wetenschappen in Nijmegen en is daar onder andere betrokken bij de Landelijke Evaluatie Onderwijsvoorrangsbeleid.

Correspondentieadres: ITS, Postbus 9048, 6500 KJ Nijmegen

Abstract

The achievements of the OVB-target groups: comparison of the results of 1988 and 1990

P. Tesser & H. Vierke. Pedagogische Studiën, 1993, 70, 122-134.

On behalf of the National Evaluation Survey on Educational Priority Policy (OVB), two institutions for educational research in The Netherlands are carrying out a longitudinal research into the development of school achievements of subsequent cohorts of primary school students. The first stage of the research took place in 1988, the second one in 1990. This article deals with the results of both stages. It starts with some analyses of the data from the second stage. After that it describes the cross-sectional and longitudinal comparison of the results of 1988 and 1990. The backlog of the students belonging to the Educational Priority Policy target groups appears to be in 1990 as great as in 1988. This applies in equal extent to students at all schools, independent of the amount of OVB facilities a school get.

Differentiatie en zorgverbreding in het basisonderwijs: de behandeling van een vernieuwingsonderwerp in onderzoek en beleid, 1984-1990

K. M. Stokking

Samenvatting

Differentiatie en zorgverbreding zijn vernieuwingsidealen en beleidsdoelen in het basisonderwijs. In advies- en beleids-nota's worden daarvan verschillende omschrijvingen gegeven, worden veel wensen geuit voor wat scholen en leerkrachten (zouden) moeten doen en worden veel factoren genoemd die daarbij bevorderend of belemmerend zouden zijn, maar wordt relatief weinig gesteund op beschikbare empirische onderzoeken. Deze laatste blijken onderling sterk te verschillen in operationalisering, steekproeftrekking en methoden, en ook in uitkomsten. Al met al echter lijkt de onderwijspraktijk in de periode 1984-1990 gemiddeld genomen weinig te zijn veranderd, terwijl slechts van één factor het belang consistent meerdere malen is gebleken: de keuzen die scholen c.q. leerkrachten zelf maken.

1 Inleiding; aanleiding en vraagstelling

Het beleid van de rijksoverheid ten aanzien van het basisonderwijs heeft in de jaren tachtig sterk in het teken gestaan van twee ontwikkelingen. De eerste was het in werking treden van de Wet op het Basisonderwijs (WBO). Daarin zijn algemene doelen van het basisonderwijs opgenomen, waaronder het bevorderen van de continue ontwikkeling van de leerlingen. Deze formulering verwees waarschijnlijk tijdens het maken van de wetstekst primair naar het opheffen van de externe differentiatie (naar leeftijd) tussen kleuterschool en lagere school. Het betreffende wetsartikel is sindsdien echter op grote schaal verbonden met een opdracht tot individualisering en differentiatie. De tweede ontwikkeling was de voortdurende verwijzing van leerlingen naar het speciaal onderwijs. Als reactie daarop werd een 'zorgverbredingsbeleid' ontwikkeld. In het kielzog van deze ontwikkelingen zijn vele onderzoeken uitgevoerd. Bij nadere beschouwing blijkt het beleid slechts beperkt te refereren aan gepubliceerde onderzoeksresultaten. Ook lopen onderzoeken onderling soms aanzienlijk uiteen. Om hierop meer zicht te krijgen is met subsidie van de SVO (ANGO 0335) een literatuurstudie uitgevoerd. Dit artikel doet daarvan verslag. (Zie voor het volledige verslag Stokking, 1992). De studie was gericht op drie vraagstellingen:

1 In welke mate lopen de beschikbare onderzoeksresultaten in de periode 1984-1989 ten aanzien van differentiatie en zorgverbreding in het basisonderwijs uiteen en hoe kunnen deze verschillen worden verklaard?
2 Wat zijn de correcte punt- en intervalschattingen van de aard en mate van differentiatie en zorgverbreding (en daarmee samenhangende factoren en condities) en de veranderingen daarin, in de populatie van alle Nederlandse basisscholen?
3 In hoeverre wordt in advies- en beleidsdocumenten gebruik gemaakt van de beschikbare onderzoekspublikaties c.q. in hoeverre is 'de discussie over differentiatie en zorgverbreding in het basisonderwijs' reconstrueerbaar?

2 Differentiatie en zorgverbreding

Leerlingen verschillen. Dit simpele feit betekent dat in scholen en in klassen de leerlingengroepen meer of minder heterogeen zijn. De heterogeniteit binnen een school wordt veelal beperkt doordat scholen hun leerlingen recruteren uit gezinnen die bepaalde kenmerken gemeenschappelijk hebben: woonwijk, opleidingsniveau, opvattingen, levensovertuiging.

Toch is de heterogeniteit in veel scholen zodanig, dat de verschillen tussen de leerlingen niet kunnen worden genegeerd. Daarmee moet dan rekening worden gehouden. Eén reden is dat het managen van en onderwijzen aan heterogene groepen moeilijk is. Dit leidt tot maatregelen die de groepen minder heterogeen maken. De bekendste is het indelen in jaarklassen: groepen die relatief homogeen zijn qua leeftijd. Een daaraan gerelateerde andere veelgebruikte maatregel is het zittenblijven: leerlingen die onvoldoende kunnen meekomen met de rest van de klas doen de betreffende jaarklas nogmaals. Deze eerste reden voor het rekening houden met de verschillen tussen leerlingen komt voort uit het gezichtspunt van de school resp. de leraren: de leerlingengroepen moeten hanteerbaar blijven.

Een tweede reden komt voort uit de leerlingen: hun mogelijkheden, behoeften, voorkeuren en wensen lopen dermate uiteen dat zij niet optimaal kunnen profiteren van het geboden onderwijs als dit niet enigszins daaraan wordt aangepast. Deze reden verwijst niet naar het hanteerbaar en dus beperkt houden van verschillen maar naar het inspelen op die verschillen. Niet de klas maar de individuele leerlingen staan centraal, de individuele ontwikkeling van elke leerling.

Deze redenen kunnen leiden tot twee duidelijk verschillende strategieën: een 'defensieve' of 'passieve' strategie, waarin verschillen worden beperkt, en een 'offensieve' of 'actieve' strategie, waarin verschillen worden opgespoord en gehonoreerd. Bij beide strategieën kunnen zich grensgevallen voordoen. Als in de passieve strategie blijkt dat een leerling 'niet langer te handhaven is', kan worden besloten die leerling te verwijzen naar een school voor speciaal onderwijs. Als in de actieve strategie blijkt dat leerlingen heel eigen mogelijkheden, behoeften en wensen hebben, kan het onderwijs min of meer worden geïndividualiseerd. In beide gevallen wordt de band met de klasgenootjes meer of minder losgelaten.

Pedagogen en didactici gebruiken voor 'het rekening houden met verschillen tussen leerlingen' van oudsher de term 'differentiatie'. Dit is een verzamelbegrip voor allerlei maatregelen en activiteiten. Bij onderwijs dat zoveel mogelijk is gericht op de individuele ontwikkeling van elke leerling afzonderlijk wordt vaak gesproken van 'individualisering'. Dit heeft dan vaak de (bij)betekenis van een pedagogisch ideaal: volledige ontplooiing van de eigen persoonlijkheid. Bepaalde groepen 'traditionele vernieuwingsscholen' hebben een dergelijk ideaal hoog in het vaandel. Het bestaan van zulke scholen is een vorm van externe differentiatie; ouders kunnen er bewust voor kiezen hun kinderen naar zo'n school te sturen.

In veel basisscholen heeft het geven van extra aandacht aan individuele leerlingen niet direct de betekenis van het bijdragen aan individuele ontplooiing zoals zojuist beschreven. Niet zozeer bijzondere talenten van de leerling maar ernstige leerproblemen vormen dan de aanleiding. Ook hier is sprake van een pedagogisch ideaal, maar dat heeft nu betrekking op het als leraar niet willen loslaten van een kind met problemen, het zich ervoor inzetten dat de leerling toch zoveel mogelijk de (instrumentele) kennis en vaardigheden onder de knie krijgt die voor het verdere leren en voor een bevredigend functioneren in de samenleving belangrijk zijn. Rond 1980 is hiervoor de term 'zorg' in zwang gekomen. De mate waarin een school in staat is om hulp te bieden aan leerlingen met leerproblemen is sindsdien ook wel de 'zorgbreedte' van die school genoemd, en het vergroten daarvan is aangeduid met 'zorgverbreding'.

De aanleiding voor deze studie was tweeledig (zie par. 1). Ten eerste leek de eerste officiële evaluatie van de WBO door het Ministerie van O&W (1988) onvoldoende systematisch de beschikbare onderzoeken te gebruiken. Ten tweede leidde een eerste vergelijking van onderzoekspublikaties tot de constatering van een aantal onduidelijkheden. Er bestaan diverse *opvattingen* over de betekenis van differentiatie en zorgverbreding, de invoeringsmogelijkheden en de te bereiken doelen (ARBO, 1984a; Span, Abbring & Meijer, 1985; Inspectiewerkgroep, 1985; Reezigt, Van Dijk & Bosveld, 1986; Wolbert, Schaap-Hummel & Span, 1986; Stokking & Leenders, 1988; Reezigt & Weide, 1989). Over de *activiteiten en maatregelen* bij differentiatie en zorgverbreding lopen de onderzoeksgegevens opvallend uiteen. Voor 'niveaudifferentiatie' bijvoorbeeld komen Reezigt e.a. (1986) op 28% (met "leerstof in verschillende niveaus") en 46% (met "ni-

veaugroepen met doorbroken jaarklassenverband"), Wolbert e.a. (1986) op 53% ("niveaudifferentiatie", op schoolniveau) en 73% ("werken met niveaugroepen", in de groepen 3 en 4) en Span e.a. (1985) op 28% (voor "regelmatig homogeen groeperen"). Voor 'tempodifferentiatie' komen Reezigt e.a. op 23% ("ongelijk tempo wordt toegelaten") en 70% ("tempoverschillen" totaal) en Wolbert e.a. op 58% ("tempodifferentiatie"). En ook worden diverse *factoren en condities* genoemd als bevorderend of belemmerend (zie over klassegrootte: Meijer, 1988; Span e.a., 1985; Reezigt e.a., 1986; Wolbert e.a., 1986; combinatieklassen: Span e.a., 1985; Reezigt e.a., 1986; Wolbert e.a., 1986; Veenman & Lem, 1985; nascholing: Meijer, 1988; Reezigt e.a., 1986; Span e.a., 1985).

3 Meta-analyse en evaluatie van onderzoeken

Aan het combineren van uitkomsten van onderzoeken is al vaker aandacht besteed (zie bijv. Churchman, 1957; Meehl, 1970; De Groot, 1969, 1977; Van Schilfgaarde, 1970; Glaser & Strauss, 1974; Stokking, 1979, 1984; Gifi, 1981; Cronbach, 1982). Als een onderzoek nog gaande is kunnen sequentiële technieken worden gebruikt. Bij reeds uitgevoerde onderzoeken en bij kruisvalidatie of replicatie blijkt vaak hoe steekproefafhankelijk onderzoeksuitkomsten kunnen zijn. Toetsing op statistische significantie beoogt de toevallige van de niet meer als toevallig te beschouwen uitkomsten te scheiden. Maar als een deel van de onderzoeken een significante samenhang tussen variabelen laat zien en een deel niet, blijft de vraag welke conclusie kan worden getrokken. Er is wel geopperd, uitgaande van het falsificatieprincipe waarin alleen verwerpingen van hypothesen tellen, om alleen de negatieve uitkomsten ('geen samenhang') mee te tellen. Maar dan zou één onderzoek met negatieve uitkomst al doorslaggevend zijn, terwijl juist het uitgevoerd zijn van meer onderzoeken met verschillende uitkomsten het probleem vormt. Bovendien biedt deze suggestie geen oplossing voor niveausteekproefonderzoek (Kersten & Bethlehem, 1981).

In de jaren zeventig is in de V.S. een aanpak ontwikkeld onder de naam 'meta-analyse' (zie o.a. Glass, 1977; Glass, McGaw & Smith, 1981; Hunter, Schmidt & Jackson, 1982; Rosenthal, 1984; Slavin, 1986; Kulik & Kulik, 1989). Dit betreft het statistisch verwerken van de uitkomsten van een groot aantal onderzoeken naar hetzelfde probleem. Uitgevoerde meta-analyses betroffen meestal onderzoeken naar het effect van een bepaalde treatment; er zijn uiteenlopende maten voor effect-groottes voorgesteld.

Meta-analyse is een puur statistische oplossing voor een typisch Amerikaans overmaatprobleem. Wij hebben die aanpak niet gevolgd, want het beschikbare aantal onderzoeken was daarvoor te gering. Ook verschillen onderzoeken in kwaliteit. Glass e.a. (1981) stellen dat het er alleen om gaat of de uitkomsten verschillen en dat de samenhang tussen onderzoekskwaliteit en onderzoeksuitkomst vaak gering is. Volgens Hunter e.a. (1982) echter kan 75% van de verschillen tussen uitkomsten van onderzoeken worden toegeschreven aan methodische fouten. Wij denken, met Hunter e.a. (1982), Cronbach (1982) en Slavin (1986), dat meer aandacht nodig is voor controle op methodologische kwaliteit dan Glass e.a. stellen.

Onderzoeken zijn ongetwijfeld ongelijk van kwaliteit. Het zou naïef zijn om zonder meer te denken dat als twee onderzoeken zijn uitgevoerd en de uitkomsten verschillen, het laatst uitgevoerde onderzoek aantoont dat het eerste niet klopt, of dat het gemiddelde van beide een getrouw beeld geeft van de feitelijke situatie (omdat beide onderzoeken voor elkaars zwakheden compenseren), of dat de werkelijkheid inmiddels kennelijk is veranderd. De onderzoeken moeten elk op zich worden doorgelicht (cf. Van Schilfgaarde, 1970) en de uitkomsten moeten zonodig vergelijkbaar worden gemaakt (cf. Gifi, 1981). Bij onvoldoende betrouwbare metingen kan daarmee maar beter niets worden gedaan. Bij verschillende operationaliseringen kan de conclusie luiden dat de onderzoeken elkaar aanvullen. Bij gelijke operationaliseringen en betrouwbare metingen maar verschillende uitkomsten is sprake van steekproeffluctuatie die verklaard moet worden uit verschillen in doelpopulatie en steekproefkader en/of bepaalde fouten.

In de opzet van deze studie was ook formele

generalisatie van de uitkomsten van de onderzoeken voorzien (zie vraag 2). Voor het generaliseren van de uitkomsten naar de populatie zijn technieken beschikbaar (Kish, 1965; Moors & Muilwijk, 1975). Ook is bekend welke fouten in steekproefonderzoek kunnen optreden en wat daarvan de gevolgen kunnen zijn (Kersten & Bethlehem, 1981; Bethlehem & Kersten, 1986). Naast de reeds genoemde meetfouten gaat het om dekkings- en kader-fouten, steekproeffouten en niet-representatieve non-respons. Gebruik van bestanden van scholen voor het trekken van steekproeven leert dat dekkings- en kaderfouten een reëel probleem kunnen vormen. Dekkingsfouten treden bijv. op doordat scholen worden gesticht en opgeheven. Van kaderfouten is sprake als adressen niet (meer) kloppen of scholen dubbel in het bestand zitten. Steekproeffouten betreffen de met aselecte trekking gepaard gaande toevalsfluctuaties. Ten aanzien van non-respons ten slotte wordt meestal gerapporteerd over het percentage algehele non-respons en worden vaak wel indicaties gegeven van de representativiteit van de respons. Voor partiële non-respons (het slechts beantwoorden van een deel van de vragen) is echter minder aandacht, en formele schattingen worden weinig gemaakt. Mogelijke gevolgen van de genoemde fouten zijn vertekening (over- en onderschatting) en/of gebrek aan precisie (brede betrouwbaarheidsintervallen). Hiervoor kan gecorrigeerd worden, mede afhankelijk van de aard van de steekproef en de benadering daarvan (adressering, rappelering), en de informatie die beschikbaar is over respondenten en non-respondenten (zie bijv. Bethlehem & Kersten, 1986).

Tijdens de uitvoering van de studie bleek dat de beoogde statistische generalisatie niet haalbaar was (zie par. 5.1).

4 De gevolgde werkwijze

4.1 Selectie van onderzoeken en beleidsdocumenten

Het domein van de studie is nader afgebakend tot rapporten en tijdschriftartikelen over onderzoeken naar differentiatie en zorgverbreding in het basisonderwijs, verschenen in de periode 1984-1990, van wetenschappelijke aard (systematisch, controleerbaar), met een survey-karakter en/of een generaliserende pretentie, en nota's van de ARBO en het Ministerie van O&W waarin naar zulke onderzoeken wordt verwezen.

Een aantal publikaties dat hieraan voldoet was reeds beschikbaar. Via een aantal kanalen zijn meer publikaties gezocht. Na een halfjaar is dit zoekproces stopgezet. Er waren toen 70 publikaties beschikbaar van 25 onderzoeken, zo'n 20 advies- en beleidsdocumenten, ongeveer 10 publikaties waarin wetenschappers een stand van zaken geven (gebaseerd op meerdere onderzoeken) en ongeveer 20 artikelen met reacties van onderzoekers of ondersteuners op onderzoeken of nota's.

Figuur 1 bevat een overzicht van de 25 onderzoeken. De onderzoeken naar differentiatie zijn genummerd van 1-12 en die naar zorgverbreding van 1-13; daarnaar wordt verwezen door de codes D1-D12 resp. Z1-Z13. In de literatuurlijst zijn de advies- en beleidsnota's opgenomen en van elk onderzoek één of enkele publikatie(s).

De verrichte studie was leerzaam ten aanzien van de mogelijkheden om op basis van publikaties van onderzoeken op een bepaald terrein tot een stand van zaken overzicht te komen. De voornaamste problemen waren de volgende. Het blijkt niet eenvoudig te zijn om van onderzoeken die meer dan enkele jaren geleden zijn afgerond de rapporten te verkrijgen; sommige publikaties blijken niet (meer) te vinden. Het komt regelmatig voor (ook in wetenschappelijke publikaties) dat wordt verwezen naar (resultaten van) onderzoek, zonder dat wordt aangegeven op welke bronnen men zich baseert. In de wetenschappelijke publikaties, zelfs als het eindrapporten betreft, ontbreken regelmatig de nodige technische gegevens en preciese kwantitatieve uitkomstgegevens. Veel onderzoeken komen tot verschillende uitkomsten omdat verschillende zaken zijn onderzocht (het gaat wel steeds over 'differentiatie' of 'zorgverbreding', maar deze zijn uiteenlopend geoperationaliseerd en soms ook gedefinieerd). Onderzoekers nemen soms in artikelen gegevens en conclusies op die niet in de (eind)rapporten staan. Een deel van de onderzoeken bestond uit een aantal deelonderzoeken, waarover ook apart is gerapporteerd. En ten slotte zijn de publikaties over een onder-

ONDERZOEK (in volgorde van jaar van dataverzameling)		INSTITUUT	N GELEZEN PUBLIKATIES	PERIODE VAN RAPPORTAGE
DIFFERENTIATIE				
1	Achterblijvers in het innovatieproces basisschool	RION	3	1983-1984
2	Problemen bij methodengebruik	VOU	3	1986-1987
3	Onderwijs in combinatieklassen	ITS	5	1985-1988
4	Leerlingbeeld en onderwijsaanbod op de basisschool	ITS	2	1985
5	Evaluatieonderzoek verspreiding GEON-resultaten	VOU	2	1986
6	Individualisering en diff. in het basisonderwijs	RION	5	1986-1988
7	Individualisering en diff. in de basisschool	VOU	3	1986
8	Effecten van differentiatie	RION	9	1988-1990
9	Periodieke peiling stand van zaken basisscholen	RION	3	1987-1988
10	Informatieverwerking en diff. op basisscholen	VOU/ISOR	1	1990
11	Knelpunten in het basisonderwijs	SCO	1	1990
12	Regionale verschillen in het basisonderwijs	VOU/ISOR	1	1990-1991
ZORGVERBREDING				
1	Determinanten van verwijzing naar speciaal ond.	RION	1	1988
2	Evaluatie ontwikkelingsprojecten bao-buo	RION	3	1984
3	Preventie van leesmoeilijkheden	KUN	1	1984
4	Evaluatie zorgverbredingsprojecten	RION	3	1986-1987
5	Schoolbegeleiding en zorgverbreding	Inspectie	1	1985
6	Verloop nascholingscursus 'speerpunt lezen'	Inspectie	1	1986
7	Voorbereidend en aanvankelijk lezen	Inspectie	1	1987
8	Toelatingsprocedures voor het speciaal ond.	RION	4	1988-1990
9	De opvang van lln. met problemen i.h. regulier ond.	RION	2	1985-1987
10	De kwaliteit v.d. onderw. zorg in reg. en spec. ond.	RION	2	1987-1988
11	Leerlingverwijzingen van basisond. naar spec. ond.	RVB	3	1986-1988
12	Evaluatieonderzoek nascholing zorgverbreding lezen	VOU/ISOR	5	1987-1990
13	Landelijk onderzoek leesvaardigheid	CITO	5	1985-1990

Figuur 1. De geanalyseerde onderzoeken

zoek soms over een aantal jaren verspreid (en kunnen dus ook op diverse tijdstippen doorwerken in dan actuele advies- en beleidsnota's).

4.2 Beschrijvingscategorieën en wijze van verwerking

Van elke publikatie is een samenvatting gemaakt. Daarbij is een lijst met beschrijvingscategorieën gebruikt (Figuur 2); de categorieën 2 t/m 8 waren alleen van toepassing bij publikaties over onderzoek. Voor de verdere verwerking zijn alle samenvattingen ingevoerd. Dit leidde tot een bestand van 800 records. Dit bestand is daarna op diverse wijzen gesorteerd.

Omdat over één onderzoek meestal meerdere publikaties zijn verschenen en deze vaak elkaar aanvullende gegevens bevatten is als eenheid de afzonderlijke publikatie genomen. Uiteraard is er vervolgens wel steeds rekening mee gehouden als meerdere publikaties betrekking hadden op hetzelfde onderzoek. Voor de beschrijving van de operationaliseringen, de onderzoeksmethoden, de respons en de uitkomsten van de uitgevoerde analyses, is de informatie zoveel mogelijk per onderzoek gecombineerd.

5 Differentiatie en zorgverbreding in onderzoek

5.1 De aard en kwaliteit van de onderzoeken

De meeste onderzoeken hadden een surveykarakter, zoals bij de selectie van de onderzoeken ook was beoogd (17 van de 25). De overige hadden wel een generaliserende pretentie maar bleken bij nader inzien toch meer het karakter te hebben van gevalsstudies (hetgeen aanvankelijk niet voldoende duidelijk was). Van de 25 onderzoeken zijn 19 door de SVO gesubsidieerd.

5.1.1 Conceptualisering

Voor vrijwel alle auteurs betekent *differentiatie* het aanbrengen van verschillen in het onderwijs aan een groep leerlingen ten aanzien van één of meer componenten: doel, leerstofaanbod, volgorde, tempo, instructie, verwerking, toetsing, beoordeling, begeleiding. Het gaat in

1. Conceptualisering van differentiatie en zorgverbreding
   - a 1. Opvattingen over differentiatie
     2. Activiteiten en maatregelen met het oog op differentiatie (school/klasniveau)
     3. Factoren en condities, van invloed op differentiatie
   - b 1. Opvattingen over zorgverbreding
     2. Activiteiten en maatregelen met het oog op zorgverbreding (school/klasniveau)
     3. Factoren en condities, van invloed op zorgverbreding
   - c Relatie differentiatie-zorgverbreding
2. Beïnstrumentering; wijze van bevraging en type onderzoek
3. Operationaliseringen; inhoud en formulering van vragen en items
4. Steekproeftrekking
   Doelpopulatie, Steekproefkader, Trekkingseenheid, Steekproeftrekking, -omvang
   Waarnemingseenheid
5. Gegevensverzameling
   Tijdstip en uitvoering van de enquête(s) en de evt. rappel(s)
   Percentages respons, weigeringen, dekkingsfouten, kaderfouten
   Percentages non-respons (= geen reactie), partiële respons
   Andere indicaties voor representativiteit
   Conclusies van de onderzoeker(s)/auteurs zelf over generalisatiemogelijkheden
6. Betrouwbaarheid
7. Validiteit
8. Analyses
   Analyse-eenhe(i)d(en), analyse-techniek(en)
   Vergelijking van instrumentarium en responsgegevens met gerapporteerde onderwerpen en aantallen respondenten, gebruikte analyse-technieken, evt. formele schattingen.
9. Uitkomsten van het bestudeerde onderzoek / uitspraken in de bestudeerde publikatie
   - a 1. Opvattingen over differentiatie
     2. Activiteiten en maatregelen met het oog op differentiatie (school/klasniveau)
     3. Factoren en condities, van invloed op differentiatie
     4. Veranderingen 1984-1990
     5. Conclusies en aanbevelingen
   - b 1. Opvattingen over zorgverbreding
     2. Activiteiten en maatregelen met het oog op zorgverbreding (school/klasniveau)
     3. Factoren en condities, van invloed op zorgverbreding
     4. Veranderingen 1984-1990
     5. Conclusies en aanbevelingen
   - c Relatie differentiatie-zorgverbreding

Figuur 2. De gehanteerde beschrijvingscategorieën

de bestudeerde publikaties vrijwel uitsluitend om differentiatie binnen scholen en binnen klassen (interne differentiatie). Voor (interne) differentiatie komen, opgesomd, de volgende intenties naar voren:

a recht doen aan c.q. rekening houden met verschillen tussen leerlingen
b recht doen aan c.q. rekening houden met individuele verschillen tussen leerlingen
c het onderwijs afstemmen op de voortgang in de ontwikkeling van de leerlingen
d het onderwijs afstemmen op grond van inzicht in de leerprocessen bij de leerlingen
e doorbreken van de uniforme leersituatie
f aanbrengen van variatie
g doorbreken van het leerstofjaarklassensysteem
h opheffen van het zittenblijven
i bevorderen van de continue ontwikkeling van leerlingen, van continue leerwegen, van een ononderbroken ontwikkeling

De genoemde componenten hebben alle betrekking op de inrichting van het onderwijsleerproces. Zij verwijzen naar meer of minder structurele of organisatorische maatregelen. Differentiëren in de zin van afstemmen van onderwijs op (individuele) leerprocessen verwijst meer naar de wijze waarop leerkrachten van moment tot moment inspelen op de reacties van de leerlingen. Dit ligt dicht aan tegen het 'diagnostiserend onderwijzen'. In de literatuur wordt het model van (leer)procesgerichte differentiatie wel genoemd, maar dit is weinig onderzocht.

Een laatste component die meestal ook wordt genoemd, betreft de beginsituatie van de leerlingen, bij het begin van een nieuw stuk onderwijs. We hebben deze niet in het rijtje hierboven opgenomen, omdat het daarbij vooral

gaat om het onderscheiden van leerlingen, en nog niet om het differentiëren van onderwijs. De beginsituatie neemt dus een andere plaats in dan de overige componenten: het zijn de verschillen tussen de leerlingen op basis waarvan beslissingen (kunnen) worden genomen over aan te brengen verschillen in het onderwijs.

De door enkele auteurs genoemde zgn. 'groeperingsvormen' (in eenvoudige gedaante: klassikaal, groepsgewijs of individueel werken) worden soms eveneens als component genoemd, soms apart opgevoerd (verschillende groeperingsvormen als wenselijkheid op zichzelf). Wij zien groeperingsvormen als organisatorische consequenties van beslissingen die ten aanzien van de eerdergenoemde componenten worden genomen, en niet als component op zich. Rekening houden met verschillen tussen leerlingen kan meer (individueel werken) of minder ver gaan (in groepjes).

Hoewel de genoemde onderwijscomponenten en ook de groeperingsvorm kunnen worden gezien als continue variabelen, is in de praktijk meestal sprake van bepaalde waarden op die variabelen. Anders gezegd: leerkrachten maken bepaalde keuzen. Zo worden bijvoorbeeld bij de component beoordeling twee of hooguit drie niveaus onderscheiden en gaat het bij het werken in groepjes meestal om groepjes van 3-5 leerlingen.

Hoewel beslissingen in principe per component kunnen worden genomen, is in de praktijk vrij vaak sprake van bepaalde combinaties. Het is gebruikelijk om hierbij te spreken van 'differentiatiemodellen' of 'onderwijsarrangementen'. In de publikaties werden de volgende modellen aangetroffen: leerstofjaarklassensysteem, werken in losser klasseverband, basisstof-herhalingsstof-extrastof/verrijkingsstof (BHE/BHV), beheersingsleren (mastery learning), werken in niveaugroepen (al dan niet klasdoorbrekend), vrije tempo werkwijze (tempodifferentiatie), differentiatie naar interesse.

Differentiatie is potentieel controversieel als het differentiatie naar doelstellingen (en leerstofinhouden) betreft: gaat het om het vergroten van de mogelijkheden van alle leerlingen om gemeenschappelijke doelen te bereiken of om het vergroten van de mogelijkheden van iedere leerling om zoveel mogelijk (differentiële) doelen te bereiken als de eigen mogelijkheden toelaten? Veel auteurs laten dit onbesproken.

Hoewel over het doel van differentiatie dus duidelijk verschil van mening (mogelijk) is, bestaat over de wenselijkheid van differentiatie echter vrijwel algemene consensus.

Zorgverbreding komt uit de bestudeerde onderzoekspublikaties naar voren als een begrip met de volgende aspecten:

- het gaat om een deel van de leerlingen, namelijk de leerlingen die dreigen vast te lopen in het onderwijsleerproces, de leerlingen met ernstige, hardnekkige problemen, de 'uitvallers', de leerlingen die door de leerkrachten als 'problematisch' worden ervaren, de leerlingen die het risico lopen om te worden verwezen naar het speciaal onderwijs;
- het gaat om deze leerlingen individueel (niet als groep);
- centraal staan de instrumentele vaardigheden die leerlingen nodig hebben om het 'gewone' onderwijs te kunnen volgen;
- het gaat om extra zorg, om aanvullende maatregelen, dus bovenop het 'gewone' onderwijs;
- het bestaat uit specifieke activiteiten van leerkrachten: signaleren, diagnostiseren, remediëren;
- het dient te resulteren in het zoveel mogelijk voorkómen van verwijzing naar het speciaal onderwijs.

Uit de publikaties blijkt dat de auteurs niet gemakkelijk kunnen omschrijven voor welke leerlingen zorgverbreding is bedoeld en om welke onderwijsactiviteiten het zou moeten gaan.

5.1.2 Operationalisering, beïnstrumentering, meting

Ook de operationaliseringen liepen sterk uiteen. In de (10) survey-onderzoeken met de meest uitgewerkte operationaliseringen (D1, D6 t/m D10, D12, Z9, Z10, Z12) bleken deze zelfs op essentiële punten te verschillen: wel of niet uitgaan van a priori opvattingen over de 'mate van' (de betreffende innovatie); wel of niet uitgaan van a priori opgestelde typen of modellen (van differentiatie); het te meten begrip (bijv. 'individualisering en differentiatie') zelf in de items opnemen of niet; de didactische

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12	Z13
schriftelijke vragenlijsten	x	x			x	x	x	x	x	x	x	x	x	x		x	x			x	x	x		x	x
mondeling interviews	x	x	x	x	x	x	x	x	x	x				x		x			x		x	x	x	x	
'bezoeken' en 'gesprekken'																x	x	x							
documentanalyse, inhoudsanalyse		x							x					x		x	x		x	x		x	x	x	
observaties, kwantitatief			x					x														x			
observaties, kwalitatief			x						x										x						
beoordelingstaak (geconstrueerde cases)				x																					
beoordelingstaak (eigen leerlingen)																						x			
logboeken		x																				x			
vragenlijsten (voor leerlingen)								x																	
stelopdrachten (voor leerlingen)			x																						
toetsen (voor leerlingen)			x					x							x							x			x
IQ-tests (voor leerlingen)								x														x			
correlaties	x							x		x			x		x										
factoranalyse				x						x					x									x	
mokkenanalyse	x								x												x				
clusteranalyse	x														x								x		
HOMALS						x		x																	
multiple regressie analyse	x																					x		x	
LISREL										x															
discriminantanalyse															x								x		
logitanalyse																							x		
t-toetsen			x																			x			x
variantieanalyse						x							x										x	x	
covariantieanalyse				x																					
multivariate variantieanalyse								x																	
VARCL								x																	
ML3												x													

Figuur 3. Gegevensverzamelingsmethoden en kwantitatieve analysemethoden in de bestudeerde onderzoeken

componenten in de vragen opnemen (per component een vraag) of in de items; gebruiken van schalen (Mokkenschaal; Likertschaal) of van losse vragen c.q. items; meten op schoolniveau en/of klasniveau; meten per schoolvak of abstraherend van schoolvak.

Bij de meeste onderzoeken zijn schriftelijke vragenlijsten en/of mondelinge interviews gebruikt; relatief weinig onderzoeken gebruikten andere typen instrumenten; met name observaties in klassen en bevraging of toetsing bij leerlingen komen weinig voor. Zie voor een overzicht Figuur 3.

Van de in totaal 17 onderzoeken met een survey-karakter wordt in de rapportage in één geval geen aandacht besteed aan de betrouwbaarheid van de gebruikte meetinstrumenten (Z13), maar kan uit het gebruik van CITO-toetsen worden afgeleid dat de betrouwbaarheid voldoende zal zijn en alleen niet expliciet is vermeld. Van de overige survey-onderzoeken besteedt een aantal in de rapportage geen aandacht aan de meetbetrouwbaarheid van de gebruikte vragenlijsten en/of kan uit de rapportage worden opgemaakt dat deze betrouwbaarheid niet is nagegaan, resp. niet kon worden nagegaan omdat uitsluitend werd gewerkt met losse items. Daarnaast zijn er 2 onderzoeken waarbij alpha-coëficiënten of rho-waarden worden vermeld maar daaraan te lage eisen worden gesteld (D8, D9).

Al met al is in 10 (van de 17) survey-onderzoeken sprake van een vastgestelde en voldoende gebleken betrouwbaarheid van de gebruikte schalen (D1, D3, D6, D10, D12, Z1, Z3, Z9, Z10, Z12). In de rapportages over de gevalsstudies wordt veelal niets gezegd over betrouwbaarheid. Waar dat wel gebeurt, is deze volgens de onderzoekers zelf in orde.

In de publikaties van meer dan de helft van de bestudeerde onderzoeken ontbreekt elke verwijzing naar de psychometrische validiteit van de gebruikte instrumenten. Bij een stuk of vier onderzoeken kan worden gesteld dat de onderzoekers zich het belang van dit aspect realiseren maar er in het onderzoek feitelijk weinig aandacht aan besteedden. In de rapportage over de resterende 7 onderzoeken (D1, D6, D8, D9, Z7, Z10, Z13) blijkt dat de validiteit van de gebruikte instrumenten min of meer systematisch is nagegaan, en wel steeds via bepaling van een of meer typen soortgenootvaliditeit. In 5 van deze gevallen bleek deze voldoende.

5.1.3 Steekproeftrekking

De schriftelijke survey-onderzoeken werkten vaak met aselect getrokken steekproeven basisscholen met een steekproefomvang tussen 200 en 1000 (gemiddeld: 670). De feitelijke trekkingswijze liep echter nogal uiteen. Enkelvoudig aselecte trekking uit het totale basisscholenbestand kwam elf keer voor. Driemaal vond stratificatie plaats, op diverse variabelen (D4, D7, D12). Het kwam vier keer voor dat de trekking plaatsvond binnen bepaalde eerst gekozen regio's, dus een getrapt aselecte trekking (D4, D5, D12, Z7). In 2 gevallen werd een (deel van een) respondentengroep van een vorig onderzoek gebruikt (D9, Z1).

Regelmatig werd aan een grootschalig schriftelijk survey een interviewronde gekoppeld, waarvoor een middelgroot aantal scholen werd geselecteerd (variërend van 36 tot 80), hetzij uit de voor het survey getrokken steekproef, hetzij uit de op het survey responderende groep, hetzij onafhankelijk. Deze interviewgroep werd soms gestratificeerd getrokken (D5).

De gevalsstudies betroffen kleinere aantallen scholen (van 5 tot 24) en deze waren meestal niet aselect gekozen.

In enkele gevallen wordt over de selectie van de in het onderzoek betrokken scholen geen nadere informatie verstrekt (D11, Z6, Z8).

Binnen de scholen waren de vragenlijsten bestemd voor resp. werden de interviews afgenomen bij de directeur en/of een of meerdere (variërend van twee tot alle) leerkrachten. De keuze van de leerkrachten werd soms overgelaten aan de directeur (D4), meestal echter werd de gewenste 'bouw' of de gewenste verdeling over de 'bouwen' aangegeven, soms werden specifieke groepen geselecteerd (variërend van 1 tot 8). Soms wordt over de selectie van de in het onderzoek betrokken leerkrachten geen nadere informatie verstrekt (Z7).

Binnen groepen werden in een 7-tal onderzoeken ook leerlingen geselecteerd, aselect dan wel gestratificeerd (op uiteenlopende variabelen), ten behoeve van de verzameling van gegevens over deze leerlingen (D4, D8, Z3, Z8 t/m Z11).

Al met al is het, afgaande op de steekproeftrekking, bij amper de helft van de bestudeerde onderzoeken mogelijk om de uitkomsten bij variabelen op (basis)schoolniveau statistisch te generaliseren. De generalisatiemogelijkheden bij variabelen op leerkracht/klasniveau en op leerlingniveau zijn nog geringer.

5.1.4 Respons, verdere reducties, generaliseerbaarheid

De gegevens van de bestudeerde onderzoeken zijn verzameld tussen april 1983 en september 1989. In enkele gevallen valt uit de publikaties niet op te maken wanneer de gegevens zijn verzameld (D2, D7, D11).

De respons in de survey-onderzoeken varieerde van 35% tot 72%, met een gemiddelde van 50%. (Op het niveau van de leerkrachten is het responspercentage op basis van de beschikbare gegevens niet goed uit te rekenen, maar het ligt waarschijnlijk wat lager). Het responspercentage neemt over de jaren niet af, dus van een eventuele dalende bereidheid tot medewerking aan onderzoek blijkt bij deze onderzoeken niets.

In diverse onderzoeken was sprake van een nadere reductie van het databestand, ten behoeve van de uit te voeren analyses, op basis van volledigheid van gegevens of van passing binnen a priori geconstrueerde categorieën. Daardoor bleef soms zelfs minder dan de helft van de respondenten over (D8).

In enkele gevallen blijft ook bij herhaalde lezing van de rapportage de (berekening van de) omvang van de respons onduidelijk (D1, Z7). Hetzelfde geldt voor de mate waarin sprake is van partiële non-respons (D6, D8). In ongeveer de helft van de survey-onderzoeken is aan de representativiteit van de verkregen gegevens min of meer uitgebreid aandacht geschonken, waarvan in vier onderzoeken (mede) op basis van een serieus non-respons onderzoek (D1, D5, D10, D12). Het komt ook voor dat onderzoekers aan de representativiteit geen enkele aandacht schenken.

Deze bevindingen betekenen dat de mogelijkheden tot statistische generalisatie nog geringer zijn dan eerder is opgemerkt.

5.1.5 Analysemethoden

In alle onderzoeken waarin kwantitatieve gegevens zijn verzameld, zijn rechte tellingen gebruikt. Vaak zijn ook kruistabellen gemaakt en/of worden frequentiegegevens gepresenteerd voor diverse subgroepen. Formele (chi-kwadraat-)toetsingen van kruistabellen kwamen amper voor. Overigens werd echter juist wel vaak afgegaan op het 'statistisch significant' zijn van de analyse-uitkomsten. Het totale aantal in de onderzoeken gehanteerde kwantitatieve technieken is groot (zie Figuur 3).

Veel publikaties zijn weinig expliciet over de aggregatieniveaus waarop de onderzoeksvariabelen betrekking hebben.

5.1.6 Rapportage

De kwaliteit, volledigheid en leesbaarheid van de rapporten lopen uiteen. Het komt regelmatig voor dat essentiële technische gegevens niet zijn vermeld (zoals de wijze van steekproeftrekking, de meetmomenten, de meetbetrouwbaarheid). Het komt ook voor, niet alleen in artikelen maar ook in eindrapporten, dat slechts samengevatte uitkomsten worden gerapporteerd.

Naast het te gemakkelijk trekken van conclusies uit data komt het ook voor dat aan formele statistische toetsingen juist teveel gewicht wordt toegekend. In beide gevallen kan zowel overinterpretatie als onderinterpretatie het gevolg zijn. Onze indruk is dat zowel bij ontbreken van toetsing als bij blindvaren op toetsing meestal sprake is van overinterpretatie.

Een opvallend verschijnsel in enkele rapporten is het soms niet goed aansluiten van conclusies op de verkregen uitkomsten. Met name komt dit voor in de vorm van het, wat we kunnen noemen, 'dichotomiseren van conclusies': de uitkomsten betreffen een glijdende schaal maar de conclusies trekken hierop een in feite willekeurige streep, waardoor uitkomsten die kwantitatief slechts gradueel of zelfs vrij weinig verschillen tot geheel verschillende bewoordingen leiden (Z8).

Een laatste gegeven is het feit dat naar aanleiding van sommige onderzoeken amper wordt gepubliceerd (bijv. Z10), terwijl andere onderzoeken leiden tot een proliferatie van artikelen (bijv. D8) (waardoor enerzijds de resultaten van het onderzoek er soms niet duidelijker op worden, anderzijds bepaalde uitkomsten in het beleid veel aandacht kunnen krijgen - verdiend of niet).

5.2 Opvattingen in de praktijk blijkens onderzoek

Over de opvattingen over differentiatie en zorgverbreding zoals die leven bij schoolleiders en leerkrachten zijn weinig gegevens aangetroffen. Uit de publikaties van Hofman, Van Aalderen en Span (1984), Inspectie (1985), Kool en Wolfgram (1986), Wolbert e.a. (1986), Stokking en Leenders (1988) en Stokking en Eberg (1991) kan men opmaken dat een klassikale aanpak (nog) veel voorstanders kent, veel leerkrachten slechts matig zijn geporteerd voor 'individualisering en differentiatie', over het belang van 'zorgverbreding' meer consensus is, men die zorgverbreding echter veelal beperkt tot leerlingen met problemen, en, ondanks die beperking, een meerderheid van de leerkrachten zegt onvoldoende tijd te hebben om aan zorgverbreding echt iets te doen.

5.3 De onderwijspraktijk blijkens onderzoek

Het is niet duidelijk in hoeverre het leerstofjaarklassensysteem (nog) bestaat. Wolbert e.a. (1986) rapporteren dat 51% van de schoolleiders aangeeft dat dit systeem bij alle vakken wordt toegepast, terwijl 63% van de leerkrachten in de groepen 3 en 4 dit aangeeft. Reezigt en Weide (1988) rapporteren voor lezen en taal over alle groepen gemiddeld percentages van ruim 30% en voor rekenen een percentage van 15%. Duidelijker blijkt uit deze onderzoeken, en dat wordt bevestigd door Van Zoelen, Mertens, Stokhuyzen en Stokking (1990) en Stokking en Eberg (1991), dat 75-80% van de leerkrachten in het basisonderwijs klassikaal instructie geeft. Op 10% van de scholen wil men de leerstof en de beheersing daarvan voor alle leerlingen gelijk houden (dit is echter niet eenduidig interpreteerbaar: is hier sprake van het leerstofjaarklassensysteem in optima forma, gaat het om beheersingsleren, en/of is de leerlingenpopulatie zeer homogeen?)

Differentiëren doen leerkrachten vooral bij de verwerking, met name naar aanleiding van verschillen tussen leerlingen in prestatie en tempo, en voornamelijk door onderscheid aan te brengen in tempo (50-60% van de leerkrachten differentieert door leerlingen zo nodig meer tijd te geven, 40-60% maakt gebruik van herhalingsstof voor zwakkere en extra stof voor snellere leerlingen (BHV-model)), en minder door onderscheid in niveau (bij taal en rekenen werkt 20-30% van de leerkrachten met niveaudifferentiatie; bij lezen ligt dit zoals bekend hoger). Slechts op zo'n 10% van de scholen is naar eigen zeggen het leerstofjaarklassensysteem helemaal losgelaten.

Apart toetsen van zwakkere leerlingen wordt in de periode 1984-1990 door constant 25-30% van de leerkrachten gedaan (Reezigt e.a. (1986) voor de situatie in 1985; Stokking & Leenders (1988) voor 1986, 1987 en 1988; Van Zoelen e.a. (1990) en Stokking & Eberg (1991) voor 1989). Activiteiten bij problemen van leerlingen met lezen zijn: even pas op de plaats laten maken van de hele klas (constant 10%), extra aandacht geven tijdens de les terwijl de rest van de klas doorwerkt (constant 80%) en aanvullende oefenstof geven (constant 50-60%). Remedial teaching wordt genoemd door constant 30% van de leerkrachten.

5.4 Bevorderende en belemmerende factoren blijkens onderzoek

In de onderzoeken waren in totaal ruim 40 mogelijk relevante factoren betrokken. Daarvan blijken alleen het onderwijsconcept van de school en de visie op individualisering en differentiatie van de leerkrachten in meer dan één onderzoek en consistent positief samen te hangen met het toepassen van differentiatie. (Differentiatie lijkt op zijn beurt overigens licht samen te hangen met zowel de prestaties van de leerlingen (licht negatief; door onderzoekers geïnterpreteerd in termen van afname van effectieve leertijd) als de verwijzingen naar het speciaal onderwijs (licht positief; te interpreteren via zowel de zojuist genoemde gemiddeld lagere prestaties als via grotere gevoeligheid voor verschillen tussen leerlingen)).

Evenals bij differentiatie lijkt ook bij zorgverbreding de visie of opvatting die de school resp. de leerkracht op het onderwijs heeft, de voornaamste predictor te zijn voor de mate waarin men zorgverbreding toepast.

6 Differentiatie en zorgverbreding in advies en beleid

Conceptualiseringen en verwachtingen

De advies- en beleidsnota's bevatten vooral uitspraken over zorgverbreding en minder over differentiatie. De omschrijving van de doel-

groep was en bleef een groot probleem: Gaat het om individuele leerlingen of om groepen leerlingen (die in dezelfde omstandigheden verkeren)? Gaat het om alle leerlingen (brede opvatting) of alleen om de leerlingen met problemen (smalle opvatting)? Gaat het om alle vakken of alleen om de instrumentele vaardigheden? De doelstelling was steeds meervoudig: zowel bevorderen van de kwaliteit van het basisonderwijs, als terugdringen van het aantal verwijzingen naar het speciaal onderwijs. Hieraan gerelateerd bleken diverse fundamenteel verschillende opvattingen te worden aangehangen, die we samenvattend kunnen aanduiden met: individualiserend, egalitair en meritocratisch. Over de wegen en middelen werd navenant verschillend gedacht, ook waar het toch vrij essentiële keuzen betreft als bijv. een accent op preventie of op remediëring. Ten aanzien van door scholen en leerkrachten te nemen maatregelen en uit te voeren activiteiten waren veel publikaties weinig terughoudend: in totaal werden meer dan 50 verschillende wensen aangetroffen.

Uitspraken over opvattingen in de praktijk
In de bestudeerde advies- en beleidsdocumenten wordt amper systematisch aandacht besteed aan opvattingen die in scholen zelf (zouden kunnen) leven over differentiatie en zorgverbreding. De betreffende documenten worden gedomineerd door omschrijvingen en wensen (zie hiervoor) en veronderstellingen over bevorderende en belemmerende condities (zie hierna).

Uitspraken over de onderwijspraktijk
In de advies- en beleidsdocumenten werden voorts vrij weinig uitspraken aangetroffen over de feitelijke gang van zaken in scholen. De ARBO liet zich hierover meestal (veel) positiever uit (zie bijv. ARBO, 1988, 1990) dan het Ministerie van O&W (zie bijv. Ministerie, 1988, 1990a, 1990b). Er werd weinig gerefereerd aan de beschikbare empirische onderzoeken.

Uitspraken over bevorderende en belemmerende factoren
In de documenten worden enkele tientallen condities genoemd. Dominerend daarbij waren in de bestudeerde periode vooral de condities met betrekking tot de leerkrachten; veel werd verwezen naar het belang van hun opleiding, motivatie, zelfvertrouwen, houding, kennis, vaardigheden, inventiviteit en inzet. Ook op dit punt werd weinig gerefereerd aan empirische onderzoeken.

7 Vergelijkingen

7.1 Vergelijking tussen categorieën publikaties

Eén van de vragen was in hoeverre in de (groepen) publikaties op andere publikaties wordt gereageerd resp. de schrijvers zich op andere publikaties baseren. Dit is relatief goed zichtbaar voor zover in beleidsnota's wordt verwezen naar ten behoeve daarvan uitgebrachte adviezen, en opiniërende bijdragen duidelijk reageren op adviezen en/of beleidsnota's. Het gebruiken of niet gebruiken van onderzoeksresultaten in adviezen en beleidsnota's is echter moeilijk hard te maken. Voor een deel ontbreekt soms gewoon onderzoek waarop men zich had kunnen baseren als in een eerder stadium onderzoek was uitgezet. Voor een ander deel is wel relevant onderzoek uitgevoerd of in uitvoering maar was er nog niet over gepubliceerd; de aansluiting tussen rapport en advies of nota is dan een kwestie van timing. Ten derde verschuiven beleidsvragen soms zodanig dat recent uitgevoerd onderzoek net niet meer de op dat moment actuele vragen beantwoordt. In hoeverre, temidden van deze modaliteiten, onderzoeksresultaten eventueel eenzijdig worden benut of zelfs genegeerd, is moeilijk vast te stellen.

Individualisering, differentiatie en zorgverbreding zijn vaak verbonden met de oproep in de WBO om te streven naar 'continue ontwikkeling' van de leerlingen. Men kan dit ook lezen als een pleidooi voor het voortdurend nagaan of leerlingen eigenlijk wel leren. Op dit punt blijkt in de bestudeerde literatuur een opvallende tegenstelling te bestaan tussen enerzijds auteurs die wijzen op de noodzaak van regelmatig toetsen van leervorderingen en anderzijds auteurs die, vanuit pedagogische overwegingen, daarvoor huiverig zijn en zelfs daartegen waarschuwen.

Maatschappelijk-politieke stellingnamen en keuzen worden door auteurs soms expliciet

gemaakt (vergelijk het onderscheid tussen een egalitair en een meritocratisch onderwijssysteem). In adviesnota's lijken zulke verschillen in opvatting echter bij voorkeur te worden toegedekt, en in het beleid leiden ze niet zelden tot verkokering (vergelijk de Onderwijsvoorrangs-ideologie met de Zorgverbredingsideologie).

7.2 Vergelijking in de tijd: verandert er iets?

Voor wat betreft de visies op de problematiek in kwestie lijkt in de bestudeerde periode weinig nieuws te zijn ingebracht. Wie het ARBO-advies over zorgverbreding uit 1984 vergelijkt met adviezen en beleidsnota's uit de jaren zeventig en met de adviezen en opiniërende bijdragen in de onderzochte periode tot 1990, ziet veelal dezelfde gezichtspunten en standpunten verwoord. Wel lijken drie aspecten de laatste jaren minder centraal te staan dan eerder: de noodzaak van voldoende geschikte materialen, het probleem van te grote klassen, en het belang van een goed functionerend schoolteam. In hoeverre dit betekent dat er nu voldoende materialen zijn, de klassen niet te groot zijn en schoolteams goed functioneren, is niet duidelijk. Wellicht zijn de scherpe kantjes van deze problemen afgesleten en worden andere aspecten nu als verhoudingsgewijs meer problematisch ervaren.

In het beleid is het accent op het belang van de bekwaamheden van leerkrachten en de noodzaak om die te bevorderen een constante gebleven. Voor het overige zijn nogal wat verschillende instrumenten en aanpakken de revue gepasseerd, en krijgen recent vooral de bestuurlijke voorwaarden voor samenwerking tussen basisonderwijs en speciaal onderwijs veel aandacht. Er is hier op twee manieren sprake van 'wijkende condities'. Ten eerste wordt de ene door het beleid benadrukte voorwaarde voor zorgverbreding opgevolgd door de andere. Ten tweede lijken de opeenvolgende voorwaarden steeds verder af te liggen van de dagelijkse interactie tussen leerkracht en leerlingen.

Men kan zich afvragen of er in de bestudeerde periode sprake is van een toename van onderzoek naar de problematiek en/of van een toenemend gebruik van onderzoeksresultaten bij het beleid. Op basis van de in deze studie verkregen bevindingen kan geen van beide vragen bevestigend worden beantwoord.

Voor wat betreft de onderwijspraktijk wordt door veel auteurs gerept over een toenemende heterogenisering van de leerlingenpopulatie, en door sommigen (ook) over toenemende verschillen tussen scholen. Op basis van deze studie kan dit niet worden bevestigd of ontkend. Wel blijkt dat in de praktijk van differentiatie en zorgverbreding in het basisonderwijs in de periode 1984-1989, in tegenstelling tot hetgeen her en der is geschreven, gemiddeld genomen géén duidelijke veranderingen optreden. De meeste percentages zijn opvallend stabiel (zie 5.3).

8 Conclusies

Ten aanzien van het bestudeerde *onderzoek* vallen vooral vier zaken op. Er worden zeer uiteenlopende operationaliseringen gebruikt, hetgeen uitkomsten moeilijk vergelijkbaar maakt. De wijzen van steekproeftrekking, de mate en aard van de non-respons en de verdere reducties en selecties maken statistische generalisatie bij veel onderzoeken in feite onmogelijk. Rapporten missen regelmatig technische gegevens waardoor onderzoeken moeilijk controleerbaar worden. En de rapportage is soms sterk samenvattend van aard en daardoor minder informatief.

Dominerende bevindingen ten aanzien van de *onderwijspraktijk* zijn: dat velen niet staan te springen om 'individualisering en differentiatie' in te voeren; dat men wel aandacht wil besteden aan leerlingen met problemen maar daarvoor weinig tijd heeft; dat veel leerkrachten het prestatieniveau van de leerlingen in hun klas zoveel mogelijk gelijk willen houden; dat de mate waarin scholen en leerkrachten feitelijk werk maken van differentiatie en zorgverbreding voornamelijk afhangt van de keuzen die ze zelf maken (de vele overige in tal van publikaties genoemde factoren en voorwaarden worden niet of niet consistent door onderzoek gestaafd); en dat in de periode '84-'90 gemiddeld genomen in de betreffende onderwijspraktijk weinig blijkt te zijn veranderd.

Ten slotte is gebleken dat onderzoeksuitkomsten niet goed worden verwerkt door het *beleid*. Er is soms meer bekend dan in beleidsdocumenten wordt voorgesteld, en ook zijn

sommige uitspraken in zulke documenten regelrecht in strijd met hetgeen uit onderzoek bekend is. In de bestudeerde beleidsnota's lijkt bovendien voornamelijk te worden verwezen naar onderzoeken waartoe het beleid zelf (in een eerder stadium) opdracht heeft gegeven. Het is onduidelijk in hoeverre hier sprake is van tekortschietende verwerkingscapaciteit dan wel dat voornamelijk bepaalde beleidsvormingsmechanismen hieraan debet zijn.

Literatuur

Aalderen, M.J. van, & Harskamp, E.G. (1987). *Periodieke peiling stand van zaken basisscholen. Deelrapport 2. Pilot-study.* Groningen: RION.

Aalderen, M.J. van, & Harskamp, E.G. (1988). *Periodieke peiling stand van zaken basisscholen. Deelrapport 3. Tussentijds verslag.* Groningen: RION.

APPO. (1984). *Ondersteuningsplan 1984-1985.* (zorgbreedte in relatie BaO-BuO). Zeist: APPO.

APPO. (1987). *Projecten regionale zorgverbredingsactiviteiten.* Zeist: APPO.

ARBO. (1984a). *Het moet ons een zorg zijn. Advies over zorgverbreding in het basisonderwijs.* Zeist: ARBO.

ARBO. (1984b). *Aanvullende adviserende notitie over zorgverbreding in het basisonderwijs.* Zeist: ARBO.

ARBO. (1986). *Advies raamplan nascholing zorgverbreding. Thema 2: begrijpend lezen.* Zeist: ARBO.

ARBO. (1988). *Voorrang aan achterstand. Advies over een integraal beleid ter voorkoming en bestrijding van onderwijsachterstand.* Zeist: ARBO.

ARBO. (1990). *Opmaat tot samenspel: advies over de binding tussen regulier en speciaal onderwijs.* Zeist: ARBO.

Bethlehem, J.G., & Kersten, H.M.P. (1986). *Werken met non-respons.* Amsterdam: CBS.

Blaauboer, S.A.A. (1984). Het innovatieproces basisschool en de plaats van de ontwikkelingsprojecten zorgbreedte, buitengewoon onderwijs-basisonderwijs daarbinnen. *Tijdschrift voor Orthopedagogiek, 23,* 138-149.

Bosveld, J.J.F., Dijk, M.H. van, Reezigt, G.J., & Weide, M.G. (1988). Differentiatiemodellen in het kleuter- en lager onderwijs. *Pedagogische Studiën, 65,* 250-259.

Churchman, C.W. (1957). *Theory of experimental inference.* New York: MacMillan.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York etc: Academic Press.

Cronbach, L.J. (1982). *Designing evaluations of educational and social programs.* San Francisco etc: Jossey-Bass.

Delsing, P.J., Jonge, J.F.M. de & Meesters, M.J. (1988). *Leerlingverwijzingen van basisonderwijs naar speciaal onderwijs. Evaluatie zorgverbredingsbeleid.* Leiden: Research voor Beleid.

Dongen, D. van (1984). *Leesmoeilijkheden. Naar diagnosticerend onderwijzen bij het leren lezen.* Tilburg: Zwijsen.

Gifi, A. (1981). *Nonlinear multivariate analysis.* Leiden: Vakgroep Datatheorie.

Glaser, B.G., & Strauss, A.L. (1974). *The discovery of grounded theory.* Chicago: Aldine.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research.* Beverly Hills/London: Sage.

Glass, G.V. (1977). Integrating findings: the meta-analysis of research. *Review of Research in Education, 5,* 351-379.

Groot, A.D. de (1977). Strategieën voor forum-convergentie. *Nederlands Tijdschrift voor de Psychologie, 32,* 397-421.

Groot, A.D. de (1969). Methodologie. Den Haag: Mouton.

Hofman, R.H., Aalderen, M.J. van, & Span, B. (1984). *Achterblijvers in het innovatieproces basisschool. Deel 2: Omvang, kenmerken en motieven.* Haren: RION.

Hunter, J.E., Schmidt, F.L., & Jackson, G.B. (1982). *Meta-analysis. Cumulating research findings across studies.* Beverly Hills/London: Sage.

Inspectie van het basisonderwijs, opleidingen en speciale diensten. (1986). *Speerpunt lezen. Inspectierapport 4.* Zoetermeer: Inspectie van het onderwijs.

Inspectiewerkgroep "Schoolbegeleiding en zorgverbreding". (1985). *Schoolbegeleiding en Zorgverbreding. Een verkennend onderzoek naar de vertaling van het zorgbreedtebeleid door de onderwijsbegeleidingsdiensten. Stand van zaken kalenderjaar 1985. Inspectierapport 1.* Rijswijk: Inspectie van het onderwijs.

Jungbluth, P. (1985). *Verborgen differentiatie. Leerlingbeeld en onderwijsaanbod op de basisschool.* Nijmegen: ITS.

Kersten, H.M.P., & Bethlehem, J.G. (1981). *Foutenbronnen in steekproefonderzoek.* Voorburg: CBS.

Kish, L. (1965). *Survey sampling.* New York etc.: John Wiley.

Kool, E., & Wolfgram, H. (1986) *Regionale samenwerkingsverbanden in het kader van zorgverbredingsinitiatieven.* Groningen: RION.

Kulik, J. A., & Kulik, C. C. (1989). Meta-analysis in education. *International Journal of Educational Research, 13,* 221-340.

Meehl, P. E. (1970). Theory testing in psychology and physics: a methodological paradox. In D. E. Morrison & R. E. Henkel. *The significance controversy* (pp. 252-266). London: Butterworth.

Meijer, C. J. W. (1988). *Verwijzing gewogen. Een studie naar de determinanten van verwijzing naar speciaal onderwijs.* Groningen: RION.

Ministerie van O&W. (1984a). *Concept-Beleidsnotitie zorgverbreding.* Den Haag, Ministerie van O&W.

Ministerie van O&W. (1984b). *Beleidsnotitie zorgverbreding.* Den Haag, Ministerie van O&W.

Ministerie van O&W. (1985). *Werken aan de basisschool, 1985-1986. Overzicht van maatregelen en aktiviteiten.* Zoetermeer: Ministerie van O&W, Ministerie van O&W.

Ministerie van O&W. (1987a). *Concept-beleidsnotitie voortgang zorgverbreding.* Zoetermeer, Ministerie van O&W.

Ministerie van O&W. (1987b). *Werken aan de basisschool, 1985-1986. Overzicht van maatregelen en aktiviteiten.* Zoetermeer: Ministerie van O&W, Ministerie van O&W.

Ministerie van O&W. (1988). *Lessen uit de praktijk. Een evaluatie van de WBO. (Concept).* Zoetermeer, Ministerie van O&W.

Ministerie van O&W. (1989a). *Concept-Beleidsnotitie Evaluatie Zorgverbredingsbeleid basisonderwijs 1985-1988, 'gevolgen voor beleid'.* Zoetermeer, Ministerie van O&W.

Ministerie van O&W. (1989b). *Concept-Beleidsnotitie Zorgverbreding voortvloeiend uit de groeibeheersing (voortgezet) speciaal onderwijs.* Zoetermeer, Ministerie van O&W.

Ministerie van O&W. (1990a). *Zo hard gelopen en toch nog zo ver te gaan: De evaluatie van de Wet op het basisonderwijs, periode 1985-1988.* Zoetermeer, Ministerie van O&W.

Ministerie van O&W. (1990b). *Weer samen naar school. Perspectief om leerlingen ook in reguliere scholen onderwijs op maat te bieden.* Zoetermeer, Ministerie van O&W.

Ministerie van WVC. (1990). *Zorg-en-de-school: een advies over de samenhang tussen regulier en speciaal onderwijs en (jeugd)hulpverlening.* Rijswijk: Distributiecentrum Overheidspublikaties.

Moors, J. J. A., Muilwijk, J. (1975). *Steekproeven.* Amsterdam, Brussel: Agon Elsevier.

Pijl, S. J., & Wolfgram, H. P. (1990). Evalueren binnen nauwe grenzen. In: M. K. van der Heijden & J. Rispens. *Zorgverbreding en afgrenzing (pp. 59-68).* Amsterdam: Swets & Zeitlinger.

Polder, K. J., Koopman, P. Lington, H. & Oud, W. (1990). *Inventarisatie van knelpunten in het basisonderwijs en van materialen voor nieuwe invoeringsprogramma's en diepteprojecten.* Amsterdam: SCO.

Reezigt, G. J., Dijk, M. H. van, & Bosveld, J. J. F. (1986). *Differentiatie op de basisschool.* Den Haag: SVO.

Reezigt, G. J., & Weide, M. G. (1988). *Differentiatiemodellen op de basisschool: Secundaire analyses stand van zaken onderzoek. Deelrapport 1.* Groningen: RION.

Reezigt, G. J., & Weide, M. G. (1989). *Effecten van differentiatie. Resultaten survey-onderzoek. Deelrapport IV.* Groningen: RION.

Reints, A. Differentiëren in en met schoolboeken. (1986). In: A. Reints & P. Span (red.). *Differentiatie in het onderwijs. Bijdragen aan de onderwijsresearch 11* (pp. 103-116). Lisse: Swets & Zeitlinger.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills/London: Sage.

Schilfgaarde, P. van (1970). *Het kennisbegrip in wetenschap en beroep.* Objectiviteit als pretentie. Alphen a/d Rijn: Samsom.

Slavin, R. E. (1986). Best-evidence synthesis: an alternative to meta-analysis and traditional reviews. *Educational Researcher, 15,* 5-11.

Span, B., Abbring I. M., & Meijer, C. J. W. (1985). *De opvang van leerlingen met problemen in het regulier onderwijs.* Groningen: SVO/RION.

Span, B. (1988). *De kwaliteit van de onderwijskundige zorg in regulier en speciaal onderwijs. Een vergelijkend onderzoek naar de onderwijskundige zorg voor kinderen met leesproblemen in basis-, LOM- en MLK-onderwijs.* Groningen: RION.

Stokking, K. M. (1979). *Toetsend onderzoek.* Groningen: Academisch Proefschrift.

Stokking, K. M. (1984). *Interpreteren en evalueren.* Deventer: Van Loghum Slaterus.

Stokking, K. M. (1992). *Wijkende condities. Differentiatie en zorgverbreding in het basisonderwijs: onderzoek en beleid 1984-1990.* Utrecht: ISOR.

Stokking, K. M., & Leenders, F. J. (1986). Effecten van een Nederlandse poging tot verspreiding van vernieuwingen. *Pedagogische Studien, 63,* 145-158.

Stokking, K. M., & Eberg, J. (1991). *Regionale verschillen in het basisonderwijs. Verslag van een onderzoek.* Utrecht: ISOR.

Stokking, K. M., & Leenders, F. J. (1988). *Effectiviteit van nascholing en nascholingsbeleid. Eindrapport van het evaluatieonderzoek naar de landelijke nascholingsoperatie speerpunt lezen.* Den Haag: SVO/ISOR.

Stokking, K. M. & Leenders, F. J. (1990). Evaluatie van een beleidsmaatregel: uitvoering en effecten van de nascholing zorgverbreding lezen. *Pedagogische Studiën, 67,* 245-260.

Veenman, S., & Lem, P. (1985). Combinatieklas vormt zware belasting voor leerkracht. Geen tijd voor zorgverbreding of differentiatie. *Didaktief, 15,* 13-15.

Veenman, S., & Lem, P. (1986). Groeperen en differentiëren in combinatieklassen. In: A. Reints & Span, P. (red.). *Differentiatie in het onderwijs. Bijdragen aan de onderwijsresearch 11* (pp. 89-102). Lisse: Swets & Zeitlinger.

Veldhuizen, G. A. van, Pijl, Y. J., Kooistra, L. & Span, B. (1983). *Achterblijvers in het innovatieproces basisschool. Deel 1: Instrumentontwikkeling.* Haren: RION.

Vinjé, M. P. (1988). *Evaluatie van twee jaar zorgverbredingsbeleid. Resultaten van het landelijk onderzoek leesvaardigheid in 1987.* Specialistisch Bulletin, nr. 66. Arnhem: CITO.

Vinjé, M. P. (1989). *Ontwikkelingen in de leesvaardigheid van leerlingen in het basisonderwijs. Resultaten van het landelijk onderzoek leesvaardigheid in 1986 en 1988.* Specialistisch Bulletin, nr. 71. Arnhem: CITO.

Weide, M. G., & Reezigt, G. J. (1988). *Differentiatie en selectie. Instrumentconstructie. Deelrapport 2.* Groningen: RION.

Weide, M. G., & Reezigt, G. J. (1990). Differentiatie en selectie in het basisonderwijs. *Pedagogische Studiën, 67,* 222-230.

Werkgroep "Rapportage Schoolwerkplanontwikkeling". (1987). *Schrijven, zeggen, doen. Een inspectieonderzoek naar de toestand van het voorbereidend en aanvankelijk lezen in scholen voor basisonderwijs. Stand van zaken 1986.* Den Haag: Ministerie van O&W.

Wolbert, R. G. M., Schaap-Hummel, I., & Span, P. (1986). *Individualisering en differentiatie in de basisschool.* Utrecht: VOU.

IJzendoorn, W. J. E. van, & Wolfgram, H. P. (1988). *Toelatingsprocedures voor het speciaal onderwijs. Deelrapport 2. Knelpunten rond het toelatingsonderzoek.* Groningen: RION.

Zoelen, L. van, Mertens, N., Stokhuyzen, Y., & Stokking, K. (1990). *Informatieverwerking en differentiatie op Nederlandse basisscholen.* Utrecht: ISOR.

Manuscript aanvaard 1-3-1993

Auteur

K. M. Stokking, senioronderzoeker en universitair hoofddocent Rijksuniversiteit Utrecht.

Adres: Rijksuniversiteit Utrecht, Vakgroep Onderwijskunde, Heidelberglaan 2, 3584 CS Utrecht.

Abstract

Differentiation and extra care in primary education: the treatment of an innovation topic in research and policy 1984-1990

K. M. Stokking. Pedagogische Studiën, 1993, 70, 135-150.

This article is based on a meta-study of 25 research projects and an evaluation of policy documents. Differentiation and extra care for pupils with problems are innovative ideals and policy goals in primary education in The Netherlands. Policy documents give different descriptions of these ideals and goals, contain a lot of prescriptions on activities by schools and teachers and assumptions about conditions, but refer relatively little to empirical research. Research projects differ quite a lot in operationalization, sampling and methods, and also in outcomes. The educational practice however did not change much (in the period 1984-1990) and only one condition became clear: the choices made by the schools and teachers themselves.

Boekbespreking

K. van den Bosch

Poor readers' decoding skills. Effects of training, task, and word characteristics

Katholieke Universiteit, Nijmegen 1991 202 pagina's
ISBN 90 9004 356 X

In het proefschrift van Van den Bosch wordt het fonologisch decoderen bij zwakke lezers op verschillende aspecten onderzocht. Algemeen wordt onder decoderen verstaan het toepassen van grafeem-foneemkennis bij het identificeren van woorden. Het wordt gezien als een fundamentele vaardigheid in de ontwikkeling van het leren lezen.

Van den Bosch opent zijn proefschrift met de vraag of een training in fonologisch decoderen, met name onder tijdsdruk, leidt tot beter lezen. Het paradigma van de training is simpel: kinderen moeten woorden en pseudowoorden hardop lezen onder wel en geen tijdsdruk. Twee varianten van tijdsdruk worden daarbij beproefd: het beperken van de presentatietijd van woorden, de zogenaamde 'flitswoord' training, en het aanmoedigen om woorden snel hardop te lezen. Het hardop lezen onder tijdsdruk doet volgens Van den Bosch een sterk beroep op het decoderen omdat via het proces van benoemen de relatie tussen orthografie en fonologie expliciet wordt gemaakt en het werken onder tijdsdruk vooral een snelle decodering vergt. Het aanbieden van pseudowoorden in een training is volgens Van den Bosch een manier om bij uitstek het fonologisch decoderen te oefenen; daar pseudowoorden niet voorzien zijn van een betekenis maar wel een legale spellingstructuur hebben kunnen ze alleen via grafeem-foneemkoppelingen gelezen worden en is er geen voordeel van bekendheid (lexicale facilitatie).

In hoofdstuk 2 worden aan de hand van twee trainingsexperimenten, waarbij gebruik gemaakt wordt van een pretest-training-posttest design, effecten van trainingen in fonologisch decoderen bestudeerd onder wel en geen tijdsdruk. De centrale vraag in het eerste trainingsexperiment is of tijdsdruk effectief is. Daartoe worden vier trainingen in fonologisch decoderen met elkaar vergeleken: 1) een training zonder tijdsdruk, 2) een training waarbij woorden voor korte tijd worden gepresenteerd, 3) een training waarbij kinderen worden aangemoedigd om snel hardop te lezen, en 4) een training waarbij een combinatie wordt gebruikt van de twee varianten van tijdsdruk. In het tweede trainingsexperiment wordt dieper ingegaan op de effecten van training 2. Centraal in dit experiment staat de vraagstelling of een training waarbij pseudowoorden voor beperkte tijd worden aangeboden effectiever is dan een training zonder tijdsdruk. Hiertoe wordt training 2 vergeleken met training 1 en met een conditie waarin geen training plaatsvindt. Alle decodeertrainingen werden gegeven aan zeer zwakke lezers, die gemiddeld tien jaar oud zijn en zo'n twee jaar leesachterstand hebben. De omvang van de trainingen bedroeg een achttiental oefensessies van elk 20 minuten.

Op basis van de resultaten uit het eerste trainingsexperiment concludeert Van den Bosch dat een training in fonologisch decoderen leidt tot een verbetering van de leessnelheid (p. 74). Tijdens de trainingen is er namelijk een sterke daling van de latentietijd die zo'n 100 ms. bedraagt bij de getrainde woorden en zo'n 300 ms. bij de getrainde pseudowoorden. Ook is er na de trainingen een verhoging van de leessnelheid op taken waarbij niet-getrainde woorden in een lijst of in een zin gelezen moeten worden. Er zijn geen verschillen in leessnelheid en accuratesse gevonden tussen de vier verschillende decodeertrainingen. Van den Bosch benadrukt het belang van zijn bevinding dat de verbetering van het lezen niet beperkt blijft tot het beter lezen van getrainde woorden maar ook generaliseert naar nieuwe woorden die niet getraind zijn.

Hoewel alle vier de decodeertrainingen tot een even grote vooruitgang leiden in leesprestatie is volgens Van den Bosch de training met flitswoorden het meest effectief. Alleen deze training bleek namelijk een positief effect te hebben op een taak die geacht werd het automatisch decoderen te meten (Plaat-Woord Interferentietaak). Kinderen die met deze variant getraind waren hadden namelijk meer tijd nodig om een plaatje te benoemen waarin als afleider een woord of een pseudowoord was geschreven dan een plaatje waarin als afleider geen woord maar een reeks medeklinkers werd gebruikt. Daaruit kan afgeleid worden dat de kinderen uit de 'flitswoord' training meer moeite hadden om de herkenning van de woorden te onderdrukken - een teken van toegenomen automatisering. Dit effect kon echter niet gerepliceerd worden in de

tweede trainingstudie zodat onduidelijk blijft of er sprake is van een betrouwbaar resultaat. De 'flitswoord' training leidt volgens Van den Bosch tot een verbetering van het accuraat decoderen (p. 75). De kinderen die in de tweede trainingstudie deze training gevolgd hadden, met alleen pseudowoorden, bleken namelijk op een leestaak in de natoets minder leesfouten te maken dan kinderen die geen training hadden gehad. Daarbij werden ook pseudowoorden die niet getraind waren nauwkeuriger gelezen. De betreffende leestaak bestond uit het hardop lezen van woorden en pseudowoorden die één voor één op de computer werden getoond. De leessnelheid op deze leestaak was echter niet verschillend van de controlegroep maar was wel hoger dan de leessnelheid die gevonden werd bij kinderen die eveneens een decodeertraining hadden gevolgd maar dan zonder tijdsdruk. Volgens Van den Bosch lokt de decodeertraining zonder tijdsdruk een tijdsconsumerende strategie uit. Daarentegen was de nauwkeurigheid van lezen bij de decodeertrainingen met en zonder tijdsdruk gelijk. Er was overigens geen effect op tekstbegrip.

Zoals gezegd, schrijft Van den Bosch de vooruitgangen in leesprestatie toe aan een verbetering van de decodeervaardigheid. Een argument dat hij hanteert is de bevinding dat de benoemsnelheid van cijfers niet vooruitgaat als gevolg van training maar die van woorden wel. De benoemsnelheid van cijfers was bij aanvang van de training al relatief hoog (590 ms.) en bleef na training op dat niveau hangen. De benoemsnelheid van woorden was daarentegen bij aanvang van de training aanzienlijk lager (1650 ms.) en was aan het eind van de training met zo'n 100 ms. gestegen. Aangezien het benoemen van cijfers geen decodering vergt - ze worden niet gelezen volgens het alfabetisch principe -, en woorden wel, concludeert Van den Bosch dat de vooruitgang in het benoemen van woorden toe te schrijven is aan een verbetering van de decodeervaardigheid. Eenzelfde conclusie trekt Van den Bosch uit de bevinding dat het benoemen van pseudowoorden sneller vooruitgaat dan het benoemen van bestaande woorden, dat overigens bij aanvang van de training al een relatief hogere startsnelheid had dan de benoeming van pseudowoorden. Pseudowoorden vergen meer decodering dan bestaande woorden omdat er geen herkenning op basis van eerdere ervaring in het spel kan zijn. Een grotere vooruitgang in het lezen van pseudowoorden zou vervolgens duiden op een verbetering van de decodeervaardigheid.

Van den Bosch vindt echter ook een resultaat waaruit opgemaakt kan worden dat het decoderen juist niet is verbeterd na de trainingen. Accuratesse en latentietijd van lange pseudowoorden bleken namelijk even snel vooruit te gaan als die van korte pseudowoorden. Volgens het dual-route model (Coltheart, 1978) zou een verbetering van het decodeerproces een groter effect teweeg moeten brengen bij lange pseudowoorden dan bij korte omdat lange pseudowoorden meer grafeem-foneem decodering vergen. Van den Bosch kon dit effect niet aantonen en geeft twee alternatieve verklaringen voor het uitblijven van een differentieel effect van woordlengte. De eerste is dat de vaardigheid van het decoderen inderdaad vooruit is gegaan maar dat het concept van decoderen zoals voorgesteld is door het dual-route model niet deugt. Het proces van decoderen hoeft namelijk niet per se het één op één *mappen* van letters op klanken in te houden maar zou kunnen opereren op multi-lettereenheden. Wanneer vervolgens in het proces van decoderen een vast aantal van die eenheden verwerkt wordt, bijvoorbeeld een tweetal eenheden die refereren naar onsets en rimes, zoals b - oom, maakt het voor systeem immers niet uit of er lange of korte woorden gedecodeerd moeten worden. Het effect van de parallelle vooruitgang zou hiermee te verklaren zijn. De tweede verklaring is dat het proces van decoderen niet is verbeterd maar een andere component in het proces van woorden benoemen wel. Een kandidaat voor deze component is het zogenaamde articulatorisch programmeren, een fase die ligt tussen het decoderen en het executeren van het spraakmotorisch programma.

In het derde en vierde hoofdstuk wordt de geldigheid van beide verklaringen onderzocht. Van den Bosch kan op grond van de resultaten van vervolgexperimenten beide verklaringen niet voldoende geldig achten. Op de eerste plaats kon hij niet aantonen dat beginnende en vaardige lezers gebruik maken van onset en rime eenheden bij het benoemen van woorden. De hypothese dat het decoderen de verwerking is van een vast aantal multi-lettereenheden is met dit resultaat niet weerlegd omdat de multi-lettereenheid alsnog van een andere aard kan zijn dan die van

de onsets en rimes. Op de tweede plaats kon Van den Bosch geen evidentie vinden dat zwakke lezers, die echter wel binnen de 'normale' range vallen, problemen hebben met het articulatorisch programmeren. Er is dan ook geen reden om aan te nemen dat deze component verbetert als gevolg van training, zegt Van den Bosch. Vervolgens beweert hij dat deze uitspraak niet te generaliseren is naar de populatie van zeer zwakke lezers in zijn trainingsstudie. Sommige onderzoekers, die Van den Bosch in zijn discussie op p. 132 noemt, konden namelijk aantonen dat er een groep van zeer zwakke lezers bestaat die niet alleen problemen heeft met het fonologisch decoderen maar waarschijnlijk ook met de articulatorische programmering. Van den Bosch acht het niet uitgesloten dat de zeer zwakke lezers uit zijn trainingsexperiment ook tot deze categorie lezers behoren.

De vervolgexperimenten geven dus geen verklaring voor het effect van de parallelle vooruitgang van lange en korte pseudowoorden, een effect dat aangeeft dat het decoderen niet is verbeterd na de trainingen. Toch spreekt Van den Bosch van een verbetering van het accuraat decoderen (p. 75), en in de laatste zin op p. 133 suggereert hij zelfs dat de snelheid van decoderen is toegenomen, vermoedelijk doelende op de toename van leessnelheid tijdens de trainingen. Volgens Van den Bosch ontbreekt alleen nog een precies plaatje van welke cognitieve processen bij het decoderen betrokken zijn. In het laatste hoofdstuk vult hij dit plaatje verder in en presenteert dan het theoretisch raamwerk waarbinnen de resultaten van zijn onderzoeken geïnterpreteerd kunnen worden. Het decoderen blijkt nu een proces te zijn dat uit verschillende fases bestaat. In de overtuiging dat de trainingen tot een vooruitgang in het decoderen hebben geleid, speculeert Van den Bosch vervolgens dat de twee varianten van tijdsdruk, die hij in zijn trainingsexperimenten beproefd had, elk een andere fase van het decoderen beïnvloeden. Zonder theoretisch te funderen waarom de ene variant, die van het limiteren van de aanbiedingstijd, aangrijpt op de beginfase van het decoderen en de andere variant, die van het benadrukken van een snelle respons, op de laatste fases van decoderen, gaat Van den Bosch verder met het vertalen van de gevonden resultaten in termen van de zo juist aangebrachte arbitraire toewijzing. Het resultaat van deze vertaling leidt op p. 139 dan ook onvermijdelijk tot een vraag die in wezen dezelfde is als de vraag die hij reeds in het begin van hoofdstuk 2 op p. 44 stelde, de vraag waarom het limiteren van de aanbiedingstijd effect heeft. Van den Bosch meent dat het antwoord in het proces van decoderen gezocht moet worden: "...een trainingsprogramma waarbij decodeervaardigheden worden geoefend door het lezen van kort aangeboden woorden en pseudowoorden kan een bijdrage leveren aan de behandeling van woordidentificatieproblemen van jonge, zwakke lezers" (p. 160).

Het aardige van het proefschrift van Van den Bosch is dat een, vermoedelijk, door de praktijk geïnspireerde vraag tot een vernuftige, met grote zorgvuldigheid voorbereide, reeks van experimenten heeft geleid. De manier waarop Van den Bosch zijn pseudowoordenlijst heeft geconstrueerd voor zijn tweede trainingstudie, opdat de gelijkenis met bestaande woorden zo min mogelijk is, is een voorbeeld van die zorgvuldigheid. De gevonden resultaten en de daaruit voortvloeiende beschouwingen zijn interessant voor de theorie van de onderliggende processen van woordidentificatie en de praktijk van behandeling van leesproblemen. Het proefschrift is helder geschreven en derhalve zeer toegankelijk voor de lezer.

Rest mij echter de vraag of de training onder tijdsdruk inderdaad tot een verbetering van het decoderen heeft geleid. Het effect van de parallelle vooruitgang blijft onverklaard en is daarmee de hardnekkigste contra-evidentie dat er een verbetering van het proces heeft plaatsgevonden, of op zijn minst gezegd niet alleen een verbetering van het decoderen. Tevens kunnen er, gezien de verschillen in startsnelheden, vraagtekens gezet worden bij de interpretatie van de resultaten bij het benoemen van cijfers en de verschillen in vooruitgang tussen woorden en pseudowoorden. De modellering van het proces van decoderen, zo laat in het proefschrift, voegt naar mijn idee weinig toe aan de interpretatie van de gevonden resultaten. Grotendeels komt dat, denk ik, doordat de voorgestelde fases van het decodeerproces geen intrinsieke relatie hebben met de wijze waarop het concept van decoderen, weliswaar iedere keer verschillend, gehanteerd wordt in de voorgaande reeks van onderzoeken. Met de gepoogde nuancering van het decodeerproces geeft Van den Bosch eerder weer dat het decoderen een vaag begrip is dan

dat het werkelijk een centrale plaats inneemt bij de gevonden effecten van de training en bij het probleem van de zeer zwakke lezers die aan de training deelnamen. Dat neemt niet weg dat er, praktisch gesproken, weer iets is toegevoegd aan onze kennis over behandeling van ernstige leesproblemen.

R. Yap

Literatuur

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151-216). London: Academic Press.

F. Goffree
Wiskunde en Didactiek 2.
Wolters-Noordhoff, Groningen 1992
292 pagina's ƒ 52,50 ISBN 90 013 4667 7

Op mijn tafel ligt een boek van Goffree (als wiskundige verbonden aan het Instituut voor Leerplanontwikkeling, als hoogleraar educatief ontwerpen verbonden aan de Universiteit van Amsterdam). Het is een bijzonder boek. In eerste instantie bedoeld voor aanstaande leraren basisonderwijs, maar eigenlijk geschikt voor iedereen die belangstelling heeft voor didactiek.

Wiskunde & Didactiek voor aanstaande leraren basisonderwijs heeft 292 pagina's. Er zijn vier secties. Sectie 1: Meten. In Sectie 2 vinden we het 'traditionele' rekenen. Sectie 3 gaat over het oplossen van problemen. Sectie 4 geeft voorbeelden voor meetkundige activiteiten. Bij elk hoofdstuk hoort een serie 'aanbevolen literatuur' voor lessen en afstudeerprojecten. Vragen en opdrachten zijn in blauw afgedrukt. De beredeneerde antwoorden en overwegingen daarbij zijn in elk hoofdstuk in een aparte afdeling (ook op blauw papier) geplaatst. Sectie 3, getiteld: 'Tussen schijn en werkelijkheid. Probleemoplossen: rijke problemen, tekst en context', boeit mij het meest. Het boek bevat verschillende ideeën voor goed onderwijs.

Wat is goed onderwijs? In goed onderwijs komen leraren, leerlingen en vakinhouden tot hun recht. Hoe creëer je goed onderwijs? Door te werken met rijke problemen! Goffree plaatst zijn studenten onder andere voor het probleem 'Graankorrels op het schaakbord'. Het is het verhaal van de man die het schaakspel ontwierp. Het wordt verteld in zes fasen. Elke fase eindigt met een vraag, die soms tot een reflectie, soms tot een bepaalde aanpak leidt. In de blauwe afdeling bij het hoofdstuk zijn de reflectieve antwoorden opgenomen.

Welke kwaliteitscriteria hanteert Goffree voor rijke problemen?

- Een rijk probleem inspireert tot vele wiskundige activiteiten. Door de open probleemstelling zijn verschillende aanpakken mogelijk. Veelal moet de leerling de wiskunde zelf creëren. Goffree noemt dat (horizontale) mathematisering. In het voorbeeld van de graankorrels betreft dit onder andere de introductie van nieuwe eenheden en het ontdekken van indirecte telmethoden.
- Een rijk probleem raakt de leerling. Dit betekent dat zijn belangstelling wordt gewekt, dat hij ervaart dat zijn kennis op dit moment te kort schiet, dat hij gemotiveerd is om actief op onderzoek uit te gaan. Bovendien is er een mogelijkheid dat de leerling een start maakt. Een rijk probleem blokkeert niet, maar daagt uit.
- In een rijk probleem speelt de context, door de tekst opgeroepen maar daarin niet vervat, een grote rol. De eigen inbreng van de student (leerling) wordt hierdoor serieus genomen, omdat de leerling door de context de mogelijkheid krijgt voor een eigen aanpak.
- De leerling krijgt de gelegenheid om onderzoekend bezig te zijn. Reflecteren over een oplossing of aanpak, van hemzelf en van anderen, wordt aangemoedigd.
- Rijke problemen geven aanleiding tot samenwerking in kleine en grote groepen. De leerlingen leren naar elkaar te luisteren en elkaar te overtuigen. Ook is het hier van belang dat de leerlingen in staat zijn constructieve voorstellen van medeleerlingen goed op te pakken.
- Een rijk probleem laat grensoverschrijdingen toe, waardoor inzicht ontstaat in de toepasbaarheid van de wiskundige activiteit.
- Een rijk probleem staat niet geïsoleerd en niet op zichzelf. Als het goed is vormt het een soort baken binnen een belangrijk leerstofgebied.

Werken met rijke problemen vereist grote deskundigheid van de leraar. Bij een rijk probleem is er namelijk geenszins sprake van een simpele probleemstelling. Neem het eerder genoemde graankorrel-probleem. Het aantal graankorrels op het vierenzestigste veld is 2^{63} en het totaal aantal graankorrels is 2^{64}-1 een getal van

20 cijfers! In de basisschool zal men daaraan niet toekomen. Dit betekent echter niet dat geen aanpak mogelijk zou zijn. Vooral in het begin zijn de verdubbelingen eenvoudig: 1 2 4 8 16 32 64 128 256 512 1024 2048 4096 8192, maar dan wordt het al snel moeilijk en moet naar vereenvoudigingen worden gezocht: 16.384, 32.768, 65.536, etc.

Het boek is op een realistische, waardevolle en ook actuele praktijk gericht: de verbeeldingskracht van de aanstaande leraar wordt er door versterkt. Voorbeeld: In een paragraaf 'weg met de breuken' komt de volgende opgave voor: Welke breuk is groter: 3/5 of 7/12?

Men kan beide breuken gelijknamig maken. Maar men kan ook de klok gebruiken; 3/5 uur is gelijk aan 36 minuten en 7/12 uur 35 minuten. Opeens zijn die onmogelijke abstracte gebroken getallen concreet geworden! Een ander aardig aspect waarop ik wil wijzen is dat het boek is gebaseerd op een congruentieprincipe: er is niet alleen sprake van een gepropageerde (lespraktijk op de basisschool), maar ook van de mogelijkheid van een geactualiseerde lespraktijk (lerarenopleiding). Als docenten van de lerarenopleiding willen dat hun studenten (aanstaande leraren) goed gaan onderwijzen, dan dienen zij zelf goed te onderwijzen.

Het boek straalt een uiterst plezierige en relevante opvatting over onderwijzen en leren en over de leraar en de leerling uit. Wiskunde en rekenen kun je leren. Wiskunde en rekenen onderwijzen kun je leren.

Zijn er dan geen nadelen? Ik signaleer een paar mogelijke problemen. (1) Serieus doorwerken van dit boek kost waarschijnlijk veel tijd. Ik vraag mij serieus af of men die tijd wel zal nemen. Men moet zich dit realiseren als men dit boek met z'n studenten gaat gebruiken. (2) De leraar-in-opleiding wordt met de grenzen van zijn kunnen geconfronteerd. De vele leeractiviteiten, de verschillende leerprocessen, de mogelijkheden tot eigen inbreng en discussie, de nieuwe inzichten, het inzetten van verworvenheden uit vorige activiteiten en de mogelijke vakgrensoverschrijdingen komen meestal pas tijdens het oplossingsproces naar voren. (3) De leraar-in-opleiding moet veel steun krijgen van de leraar-die-opleidt. Dit betekent dat de leraar-die-opleidt van te voren constructieve analyses op het arsenaal aan (rijke) problemen moet toepassen. Hiermee wordt een subjectief element ingebracht, want, wat de een direct aan oplossingsmogelijkheden ziet wordt door de ander over het hoofd gezien. Dit geldt voor de leraar-in-opleiding zowel als voor de leraar-die-opleidt.

Het is bekend dat aanstaande leraren basisonderwijs nogal eens moeite hebben met rekenen en wiskunde. Hier krijgen ze een uitstekende kans om te oefenen met rijke problemen die zij in de lespraktijk kunnen gebruiken. Het is een boek waar ik vrolijk van word.

H. Stroomberg

D.R. van Peer

Sociale ondersteuning van kinderen met leerstoornissen

Swets & Zeitlinger, Lisse 1991,
146 pagina's *f* 34,50 ISBN 90 265 1233 3

In de studie van leerproblemen is de neurologische oriëntatie van begin af aan dominant geweest. Sedert Strauss' en Lethinens conceptie van het *brain injured child* is ze bovendien als vanzelfsprekend gepresenteerd. Men spreekt dan ook liever van leer*stoornissen* dan van leer*problemen*. De empirische staving van een dergelijke oriëntatie is evenwel tot dusverre uiterst pover gebleken. 'Neurologen' zoals Swanson (1988) en Bryan, Bay en Donahue (1988) geven dat hun 'critici', waaronder prominenten als Kavale (1988) en Siegel (1988), onomwonden toe. Het vermag hen echter niet tot andere gedachten te brengen. Gelet ook op de gangbare definities van de problematiek (vgl. o.m. Dumont, 1978) blijft men er rotsvast van overtuigd dat er met de hersenen van sommige kinderen iets mis is waardoor zij in het leren van rekenen of taal ernstig worden belemmerd. Maar, waar men de zekerheid moet ontberen, wordt de zogenaamde discrepantie-formule als substituut gehanteerd. De tegenstelling tussen een relatief hoge intelligentie (lees: IQ) en de desondanks mager uitgevallen schoolprestaties heet te verwijzen naar de weliswaar nog niet aantoonbare, maar met de voortgang der wetenschap mettertijd ongetwijfeld aan het licht te brengen hersen-afwijking.

Nu gaat alle leren vanzelfsprekend via de hersenen. Maar dat wil nog niet zeggen dat wanneer het leren stokt, dit ook direct aan een dysfunctionerend brein moet worden toegeschreven. De werkelijkheid zou wel eens ingewikkelder

kunnen zijn. Ook Van Peer lijkt dit standpunt toegedaan. Dat zou althans opgemaakt kunnen worden uit de eerste twee hoofdstukken van zijn dissertatie. Zo citeert hij op pagina 6 met instemming Rispens en Van den Bos die het discrepantie-idee in relatie tot leerproblematiek van de nodige vraagtekens voorzien; discrepanties laten zich immers ook bij 'gewone' kinderen waarnemen. Een pagina verder lijkt hij, conform Stevens, moeite te hebben met het 'simplistisch unicausaliteitscriterium'; de sociale omgeving van kinderen - leerkrachten en ouders bijvoorbeeld - zou ook wel eens van invloed kunnen zijn. Maar is die ook bepalend voor de genese van leerproblematiek; voor sòmmige leerproblemen wellicht? Even lijkt Van Peer die indruk te wekken, en wanen we ons terug in de jaren zeventig en tachtig toen sociologen en sociaalpsychologen het bastion van de strikt kindgeoriënteerde psychologen en orthopedagogen trachtten te bestormen; ze moesten weliswaar het onderspit delven maar bepaald niet altijd op grond van een wetenschappelijke argumentatie. "De studies naar de relatie tussen opvoedersgedrag en prestatiegedrag bij kinderen met en zonder specifieke problemen", aldus Van Peer op pagina 11, "suggereren dat het opvoedersgedrag sterk afhankelijk is van de verwachtingen die de ouder heeft van de prestatie van het kind." En op pagina 12 lezen we de conclusies van een onderzoek van Lavelle: "Ouders van kinderen met leerstoornissen attribueerden het falen van hun kind aan stabiele oorzaken, in casu gebrek aan capaciteiten en verwachtten daarom weinig van hun intellectuele prestaties."

Kortom, Van Peer lijkt te redeneren in de lijn van Rosenthals en Jacobsons fameuze "Pygmalion in the Classroom" (1968). Zij achtten een onjuiste of door vooroordelen ingegeven definitie van de situatie de oorzaak, tenminste de aanleiding, voor een zichzelf waarmakende voorspelling; kinderen zouden met andere woorden wel eens leerproblematisch kunnen worden, zonder het aanvankelijk geweest te zijn. Nogmaals: zo lijkt Van Peer te redeneren. Maar de schijn bedriegt. Of nauwkeuriger: Van Peer hinkt op twee gedachten. Zegt hij op pagina 7 zich weinig te willen bekommeren om de 'feitelijke' oorzaak van de leerproblemen, en toont hij zich, blijkens de inleiding op hoofdstuk 3 vooral geïnteresseerd in de zogenaamde opvoedingsattitude van ouders, en in de al dan niet sociale ondersteuning verlenende interacties met hun kinderen, hij blijft uitdrukkelijk over *stoornis* praten, en - belangrijker - blijkt bij de operationalisatie van dat concept te hechten aan de hoogte van het IQ, inclusief de vermaledijde discrepantie-gedachte (vgl. p. 35 e.v.). En dat kan toch niet aan zijn promoter Rispens hebben gelegen.

Hoe schijnbaar inconsistent ook, de studie verliest daarmee natuurlijk niet meteen aan waarde. Het blijft interessant om na te gaan of ouders anderssoortige verwachtingen, en dienovereenkomstig anderssoortige interacties jegens hun kinderen ontwikkelen al naar gelang het type school dat zij bezoeken. Maakt het wat uit of kinderen in het LOM-onderwijs zijn geplaatst of dat zij zich, ondanks de geringe prestaties, in het reguliere onderwijs wisten te handhaven? Van Peer vermoedt van wel, en dat lijkt ook alleszins gerechtvaardigd. "Het is duidelijk", stelt hij, "dat de manier waarop ouders de rol van de school percipiëren in relatie met de ontwikkeling van hun kind, belangrijke gevolgen kan hebben voor hun verwachtingen en de manier waarop ze interacteren met het kind" (vgl. p. 37).

Ten behoeve van zijn empirisch onderzoek onderscheidt hij een drietal groepen ouders; moeders om precies te zijn. Vijfentwintig moeders wier zonen een LOM-school bezoeken, de LOM-groep; 25 moeders van zonen zonder leerstoornissen binnen het gewone basisonderwijs, de BAO-groep, en 25 moeders wier zonen als leergestoord aangemerkt worden, maar niettemin de reguliere school bezoeken, de BLP-groep genoemd. In drie opzichten vergelijkt hij ze: naar de opvoedingsattitude, waaronder o.m. het aspiratieniveau, de cognitieve stimulering en de onderwijsondersteunende houding worden begrepen; naar de verwachtingen die de moeders over het presteren van hun zonen hebben; en ten slotte naar het gedrag dat zich tussen moeder en zoon ontwikkelt bij de oplossing van vier aan het kind voorgelegde taken.

De BLP-groep is uiteraard de boeiendste van de drie, zeker wanneer de uitkomsten daarvan vergeleken worden met die van de LOM-groep. De attitude en de sociale ondersteuning van ouders van leerproblematische kinderen binnen het reguliere onderwijs, wordt immers niet gekleurd door de mogelijk negatieve effecten die een LOM-plaatsing met zich mee kan brengen.

Maar dan komt het er wel op aan de nodige zorgvuldigheid bij de samenstelling van de steekproef te betrachten. Welnu, daar valt het een en ander op af te dingen. Het moet gezegd: de auteur verhult dat ook niet (vgl. p. 39). Directies beweerden dat op hun basisscholen geen kinderen voorkwamen die aan de vereiste selectiecriteria voldeden, of veronderstelden dat ouders geen medewerking aan het onderzoek zouden willen verlenen et cetera. Maar misschien legde Van Peer het hoofd wat al te gauw in de schoot. In elk geval besloot hij 18 van de 25 kinderen van de BLP-groep uit twee particuliere psychologische praktijken te recruteren, daarmee de mogelijkheid tot een niet te verwaarlozen 'bias' introducerend. Want, niet uit te sluiten valt, dat ouders die, ongeacht de mening van de school, zèlf hun kind als problematisch percipiëren, een ander type opvoedingsmilieu representeren dan de ouders van de twee overige groepen kinderen. Van Peer rept met geen woord over die mogelijkheid, terwijl uit de steekproefgegevens op p. 42 blijkt dat met name het opleidingsniveau, maar tevens het beroepsniveau van de ouders van de BLP-groep beduidend hoger uitvallen dan dat van de LOM- en BAO-groep. De auteur constateert het wel, maar daar is dan ook alles mee gezegd. Des te opmerkelijker omdat hij in de discussie in hoofdstuk 7 van het boek wijst op het belang van de homogeniteit naar sociaal-economische status van steekproeven. "Wordt in onderzoek (...) voor de SES niet gecontroleerd", aldus Van Peer, "dan bestaat de kans, dat gevonden verschillen niet kunnen worden toegeschreven aan de onderzochte variabele (bijvoorbeeld de gevolgen van een leerstoornis op de opvoedingsattitude van de ouders), maar aan opvoedingsattitudes die horen bij een bepaalde SES en die relatief ongevoelig kunnen zijn voor bijzondere kenmerken van het kind" (vgl. p. 89). Of relatief *over*gevoelig, had hij daar nog aan toe kunnen voegen.

Hoe het ook zij, qua attitude kon Van Peer geen verschillen tussen de moeders van de onderscheiden groepen constateren. De schoolsituatie van de kinderen lijkt dus niet van invloed op de opvoedingshouding. Ook blijken de verwachtingen van moeders met een leergestoorde zoon over zijn presteren, om het even of hij nu de LOM- of de reguliere basisschool bezoekt, nauwelijks minder hooggespannen dan die van moeders van 'gewone' kinderen. Maar het meest relevant van alles: ook de interacties tussen moeder en kind lijken niet wezenlijk door de schoolsituatie bepaald te zijn. Dat valt althans niet op te maken uit de figuren waarin de z-waardenpatronen van de diverse onderscheiden gedragsobservatiecategorieën staan weergegeven. En daar, waar zich wel een enkel verschil aftekent, acht Van Peer ze terecht van weinig betekenis (vgl. 77-78).

Het heeft er de schijn naar dat Van Peer deze conclusies met tegenzin trekt. Het is natuurlijk ook prettiger om uitkomsten te kunnen presenteren die wat meer stroken met de theoretische uitgangspunten. Het doet dan ook wat geforceerd aan om op p. 92 te lezen dat "met enige reserve gesteld (zou) kunnen worden, dat de interacties in de BAO groep adequater zijn dan in de LOM en BLP groep." De wens zal wel de vader van de gedachte zijn geweest. Een wens, die wellicht eerder in vervulling had kunnen gaan wanneer Van Peer de interacties tussen moeder en zoon bestudeerd had aan de hand van meer 'uit het leven gegrepen' taken dan het oplossen van het Wiggly blok, de anagrammen, de ei puzzel en de woordparen. Maar toegegeven: als buitenstaander is het makkelijk praten.

J. Th. A. Bakker

Literatuur

Bryan, T., Bay, M., & Donahue, M. (1988). Implications of the learning disabilities definition for the regular education initiative. *Journal of Learning Disabilities, 21*, 23-28.

Dumont, J. J. (1978). *Leerstoornissen (1): Theorie en model*. Rotterdam: Lemniscaat.

Kavale, K. A. (1988). Epistemological relativity in learning disabilities. *Journal of Learning Disabilities, 21*, 215-218.

Rosenthal, R., & Jacobson, L. F. (1968). *Pygmalion in the classroom. Teacher expectation and pupil's intellectual development*. New York: MacMillan.

Siegel, L. S. (1988). Definitional and theoretical issues and research on learning disabilities. *Journal of Learning Disabilities, 21*, 265-266.

Swanson, H.L. (1988). Toward a metatheory of learning disabilities. *Journal of Learning Disabilities, 21*, 196-209.

G. W. Meijnen & F. S. J. Riemersma

Schoolcarrières: een klassenkwestie? De schoolloopbanen van leerlingen, bezien in relatie tot de invloeden die maatschappelijke en binnenschoolse determinanten hierop hebben

Swets & Zeitlinger, Lisse 1992
140 pagina's ƒ 31,– ISBN 90 265 1323 2

'Schoolcarrières: een klassenkwestie?' van Meijnen en Riemersma is de neerslag van een literatuurstudie die deel uitmaakt van het PSO-programma 'De schoolcarrière van kinderen, bezien in relatie tot de invloeden die maatschappelijke en binnenschoolse determinanten hierop hebben'. Dat programma wordt uitgevoerd in een samenwerking van Amsterdamse en Rotterdamse onderwijs-onderzoekers. Hoewel de vraagstelling naar haar aard een overwegend sociologische is, werken in het programma vele disciplines samen. Dat is bijna kenmerkend voor veel van de actuele onderwijsresearch, zoals ook juist die sociologische invalshoek, gaandeweg het onderzoek nog wel eens zoek wil raken.

De literatuurstudie is, zoals dat tegenwoordig gaat, vooral gebaseerd op een omvangrijke literatuursearch langs ERIC, ADION, DION en andere systemen, wat meestal leidt tot onvoorstelbaar veel kleine bijna-weetjes en weinig grote lijn. De auteurs ontkomen aan die tendens nauwelijks, maar dat was vermoedelijk ook niet hun opzet. De literatuurstudie is uitgezet ter ondersteuning van empirisch onderzoek en de auteurs zijn vooral op zoek naar nieuwe, kleine aanvullingen en aanwijzingen ter completering van de eigen designs, veel meer dan naar de grote lijnen over onderwijs en ongelijkheid.

Wie eens snel wil opzoeken, wat er aan recent empirisch onderzoek te verhapstukken valt rond het thema onderwijs en ongelijkheid, die kan met dit handzame werkje goed uit de voeten. Waarschijnlijk leidt dat in veel gevallen tot hetzelfde adres. De aangehaalde literatuur is keurig gecategoriseerd en vooral toegespitst op de voorschoolse periode en de ontwikkeling buiten de school. Op al die punten is een afzonderlijk hoofdstuk gewijd aan de problematiek van allochtone leerlingen. Hoe rijk het Nederlandse onderwijsonderzoek op dit punt is, kan de lezer nauwelijks ontgaan. Dat juist de empirische benadering in Nederland veel twijfel heeft doen ontstaan over de receptuur van een 'effectieve school', dat wordt zeker niet onder het karpet geveegd, maar juist hier zou je wat meer sociologisch framework verwachten.

In de ondertitel van het boek is sprake van 'invloeden die maatschappelijke en binnenschoolse determinanten' hebben op de schoolloopbanen van leerlingen. Natuurlijk was het boeiend geweest, als ook duidelijk werd, wat de maatschappelijke determinanten zijn van het feit dat het binnenschools gaat zoals het gaat. Daarover vermag de onderwijssociologie meer te zeggen dan je uit alleen recent empirisch onderzoek kunt afleiden. In die zin is het boekje er een duidelijk bewijs van, hoezeer die ene onderwijswetenschap verstek laat gaan, op het punt van het vasthouden van zijn eigen vraagstellingen. Meer dan driekwart van de aangehaalde literatuur past onder het 'blame the victim'-label, als lag het probleem in eerste instantie bij de ouders, terwijl de vraag naar het karakter van het standenonderwijs onder de nieuwe conditie van een multiculturele onderklasse, nauwelijks naar onderzoek wordt vertaald. Hoe het komt, dat scholen met hoge percentages kansarme leerlingen lage streefdoelen hanteren, waarom leerkrachten op zulke scholen er geen probleem van maken, als leerlingen hele leerjaren achterop raken, dat zijn vragen die ook horen bij de titel, nu niet met een vraag- maar met een uitroepteken achter het woord klassenkwestie!

P. Jungbluth

Mededelingen

75 Jaar Nutsseminarium

In 1993 vieren de Pedagogische Opleidingen van het Nutsseminarium voor Pedagogiek – tegenwoordig de Afdeling Pedagogiek van de Hogeschool van Amsterdam – het 75-jarig bestaan. Ter gelegenheid van die feestelijke herdenking wordt op 5 juni a.s. in de Aula van de Universiteit van Amsterdam een congres gehouden, tijdens welk twee boeken gepresenteerd zullen worden – één over kwaliteitsprojecten in het onderwijs, en één over actuele ontwikkelingen in de pedagogische beroepen.

Voor nadere inlichtingen: Hogeschool Amsterdam, Afdeling Pedagogiek, Herengracht 266, 1016 BV Amsterdam, tel.: 020-550 16 16.

Landelijke studiedag: Nieuwe gereedschappen voor docenten in het BVE

De Educatieve Faculteit van de Hogeschool Gelderland en BVE procescoördinatie organiseren op 10 juni 1993 te Arnhem een landelijke studiedag die als motto heeft meegekregen 'Docent Centraal'. In het ochtendprogramma worden in een aantal inleidingen verschillende dimensies van de veranderende rol van de docent in het BVE belicht. Het middagprogramma bestaat uit vier minisymposia die elk vanuit een andere invalshoek nieuwe inzichten en toepassingen aandragen voor de onderwijskundige vormgeving van het BVE.

Voor meer informatie kunt u contact opnemen met het secretariaat studiedag 10 juni, mevr. E. v.d. Heuvel, Postbus 7130, 3430 JC Nieuwegein, tel.: 03402-41731.

Stipendia voor Opleidingskundig Onderzoek

Het bestuur van de Foundation for Corporate Education – FCE – Stichting Opleidingskunde stelt in 1993 een bedrag van *f* 40.000,– ter beschikking voor onderzoek op het terrein van opleidingen en leerprocessen in arbeidsorganisaties. Het bestuur nodigt onderzoekers uit om onderzoeksvoorstellen in te dienen, die in aanmerking komen voor financiële ondersteuning in de vorm van stipendia, in het bijzonder voorstellen die zich richten op probleemstellingen rond de instrumentatie van een lerende organisatie, en de vaardigheden van managers die vormgeven aan hun opleidersrol.

Belangstellenden kunnen het stipendiumreglement opvragen bij: Foundation for Corporate Education, Stipendiumcommissie, Dronkelaarseweg 13, 3784 WB Terschuur.

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
18e jaargang, nr. 1, 1993

Bodemloosheid: de vruchtbaarheid van een (ortho)pedagogische benadering, door D. Broos en P. Ghesquière
Kinderopvang op zijn Nederlands, door B. Levering
De professionele ontwikkeling van leerkrachten gezien vanuit het spanningsveld tussen autonomie en collegialiteit, door M. Clement, K. Staessens en R. Vandenberghe
Onderwijsverandering in het Verenigd Koninkrijk in de jaren tachtig, door R. Jonathan
Instrumentele rationaliteit: ook in België?, door J.L. Vanderhoeven
Ondernemend onderwijs, door P. Baggen

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 3, 1993

Gezinsbelasting van ouders van residentieel opgenomen kinderen vergeleken met die van ambulant behandelde gezinnen, door P.M.A. Wels en L.M.H. Robbroeckx
Luisterend lezen met pauzes, door P. Reitsma
Indicatiestelling in de jeugdhulpverlening, door M. Faas

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 4, 1993

Het lezen van pseudowoorden en bestaande woorden. Deel I: De constructie en normering van de Klepel, door K.P. van den Bos, A.J.M. Scheepstra en H.C. Lutje Spelberg
Het meten van gezinscohesie: een vergelijkend onderzoek van GDS en NTG, door E.E.J. De Bruyn, J.H.L. Oud, J.M. Caris en T. Faasse

Het gaat goed met de psychodiagnostiek, of niet soms? door J. ter Laak

Ontvangen boeken

Berkel, H. J. M. van, & Bax, A. E., *Beoordelen in het onderwijs*. Bohn Stafleu Van Loghum, Houten, 1993, *f* 45,–.

Boekaerts, M., & Simons, P. R. J., *Leren en instructie. Psychologie van de leerling en het leerproces*. Dekker & Van de Vegt, Assen, 1993, *f* 69,50.

Gelderen, A. van, *De evaluatie van spreekvaardigheid in communicatieve situaties* (dissertatie). SCO-Kohnstamm Instituut, Amsterdam, 1992.

Groot, R. de, Boerema, J. H. D., & Paagman, C. J., *Bladergroen. Voorwaarden voor opvoeding en onderwijs*. Intro, Nijkerk, 1993, *f* 29,50.

Haaf, N. ten, *Opvoedingsdimensies: convergente en discriminante validiteit* (dissertatie). Katholieke Universiteit, Nijmegen, 1993.

Hermans, Chr., *Vorming en perspectief: grondslagen studie over identiteit van katholiek onderwijs*. ABKO/Gooi & Sticht, Den Haag/Baarn, 1993, *f* 39,50.

Meeus, W., & Hart, H. 't (red.), *Jongeren in Nederland*. Academische Uitgeverij, Amersfoort, 1993, *f* 37,50.

Mönks, F., & Peters, W. (Eds.), *Talent for the future*. Van Gorcum, Assen, 1992, *f* 55,–.

Pethke, H., Schiet, M., Verheul, K., *Opvoeden is maatwerk. Praktische informatie voor ouders van kinderen in de basisschoolleeftijd*. Kosmos - Z&K, Utrecht, 1992, *f* 39,90.

Rietdijk, L. E., *Teambegeleiding en individuele werkbegeleiding*. Uitgeverij SWP, Utrecht, 1993, *f* 39,–.

Roelofs, E. C., *Teamgerichte nascholing en coaching. Een experimentele studie in scholen met combinatieklassen* (dissertatie). Katholieke Universiteit, Nijmegen, 1993.

Schoorl, P. M., *Opvoeding en opvoedingsproblemen in het gezinsleven*. Swets & Zeitlinger, Lisse, 1993, *f* 45,–.

Singer, E., *Kinderopvang, goed of slecht?* Uitgeverij SWP, Utrecht, 1993, *f* 26,50.

Steutel, J. W., *Deugden en morele opvoeding. Een wijsgerig-pedagogische studie*. Boom, Meppel, 1992, *f* 28,50.

Verheij, F., *Behandelingsplanning in de jeugdhulpverlening*. Uitgeverij SWP, Utrecht, 1993, *f* 49,50.

Verhoeven, L., Extra, G., Konak, Ö. A., Narain, G., Zerrouk, R., *Tweetaligheid bij allochtone kleuters. Opzet en resultaten van een proefpeiling bij Turkse, Marokkaanse en Antilliaanse/Arubaanse kleuters*. Tilburg University Press, Tilburg, 1993, *f* 29,–.

Versloot, A. M., & Ploeg, D. A. van der, *Leerlingen: samen en apart*. Rijksuniversiteit Utrecht, 1992.

Verstegen, R., & Lodewijks, H. P. B., *Interactiewijzer. Analyse en aanpak van interactieproblemen in professionele opvoedingssituaties*. Dekker & Van de Vegt, Assen, 1993, *f* 37,50.

Walt, J. L. van der, *Fundamentele opvoedkunde en die ontisiteit van opvoeding*. RGN Uitgevers, Pretoria, 1992.

Wolfhagen, I., *Kwaliteit van klinisch onderwijs* (dissertatie). Datawyse/Universitaire Pers Maastricht, 1993.

Inhoudsopgave

Pedagogische Studiën Jaargang 70, nummer 3/1993

N. A. J. Lagerweij
Implementatie van basisvorming: de start van een dieptestudie 162

G. Theloosen en W. H. J. van Bon
Herhaling van woorden en lettergroepen: oefeneffecten bij zwakke lezers 180

J. Chr. Perrenet
De invloed van het gebruikte leerboek op het maken van specifieke fouten bij wiskundeproblemen 195

Discussie
J. Masschelein
De dood van de universiteit 206
P. Baggen en I. Weijers
'Storen' – De laatste strohalm van het idealisme 214
J. Masschelein
Het belang van het onderscheid 217

Kroniek
J. Imants
De jaarlijkse bijeenkomst van het 'International Congress for School Effectiveness and Improvement': Norrköping 1993 220

Boekbesprekingen 222

Mededelingen 236

Implementatie van basisvorming: de start van een dieptestudie

N. A. J. Lagerweij*

Samenvatting

Invoering van basisvorming betekent voor scholen in het voortgezet onderwijs een omvangrijke bundel van ingrijpende veranderingen. Het wettelijk kader geeft scholen vrijheid over diverse zaken eigen onderwijskundige en schoolorganisatorische beslissingen te nemen. In het Bavo-project worden drie brede scholengemeenschappen gedurende vijf jaren gevolgd bij hun vormgeving van de basisvorming. Deze dieptestudies moeten zicht geven op de factoren die op meso- en microniveau bijdragen aan de invoering en aan het bereiken van de doelstellingen van basisvorming.

Het onderzoek is gericht op drie clusters van vragen te weten de schoolkenmerken, het procesverloop, de vormgeving en resultaten van onderwijsleerprocessen. Voor elk cluster zijn innovatietheoretische verwachtingen opgesteld, die in de casestudies getoetst worden. Dit artikel bevat een overzicht van de gegevens uit de voorbereidingsperiode van 1991 tot 1993 betreffende de beginsituatie van de drie scholen.

1 Inleiding

In 1993 staan scholen in het Nederlandse voortgezet onderwijs voor een ingrijpende veranderingstaak. Zij moeten dan de wettelijk voorgeschreven basisvorming gaan vormgeven in de onderwijspraktijk. Invoering van deze basisvorming gaat gepaard met allerlei andere wijzigingen. Tegelijkertijd kunnen scholen betrokken zijn bij operaties betreffende schaalvergroting (fusies, e.d.) en bij invoering van nieuwe administratieve procedures rondom personeelsbeleid en financiën.

Het is duidelijk dat scholen voor voortgezet onderwijs gedurende enkele jaren een veelheid van veranderingen moeten aanbrengen of ondergaan. De invoering van basisvorming kan leiden tot wijzigingen in het functioneren van de schoolorganisatie, tot aanpassingen van het leerplan, tot verschillen in leerresultaten (verhoging peil jeugdonderwijs) en tot veranderingen in het onderwijzen. Deze veranderingen voltrekken zich niet los van elkaar, maar staan in nauwe onderlinge samenhang. Hierbij kunnen zich wat de vormgeving betreft van school tot school verschillen voordoen. De wettelijke voorschriften hebben het karakter van een kaderwet, die aan scholen enige speelruimte geeft om tot eigen invulling te komen. Een belangwekkende vraag is of en hoe de scholen erin slagen deze beleidsruimte op eigen wijze te vullen (zie ook Sleegers & Bergen, 1992).

Theorievorming over onderwijsinnovatie laat duidelijk zien dat er vele verklaringen mogelijk zijn voor het falen en slagen van vernieuwingsprojecten. Uit dit kennisbestand valt af te leiden dàt veel factoren invloed hebben op het verloop ervan (zie o.m. Lagerweij & Voogt, 1990; Hopkins, 1990; Fullan, 1991; Louis & Miles, 1990; Dalin, 1992). Het planmatig toepassen van deze innovatiekennis blijkt steeds sterk context gebonden te moeten plaats vinden. Elke innovatie (elk project) heeft eigen min of meer unieke karakteristieken, waardoor algemeen geldende regels toch weer een aanpassing ondergaan.

Onderzoekers[1] van de Vakgroep Onderwijskunde in Utrecht, werkzaam in het kader van het onderzoekprogramma 'Ontwikkeling van onderwijs en opleiding', besloten daarom in een onderzoek na te gaan welke factoren op

* Met dank aan H. Franssen, J. van der Linden, E. Roelofs, J. C. Voogt, L. Vriens, L. van Wessum en de student-assistenten M. Mettes en O. Treep. Zij werken mee aan het project en hebben daardoor en met hun commentaar op eerdere versies in belangrijke mate bijgedragen aan dit artikel.

meso- en microniveau bijdragen aan het invoeren en het bereiken van de doelstellingen van de basisvorming in scholen voor voortgezet onderwijs. Het doel van dit onderzoek is bij te dragen aan theorieontwikkeling op het gebied van onderwijsverandering en inzicht te krijgen in de wijze waarop onderwijsbeleid van de overheid kan bijdragen aan verbetering van de onderwijspraktijk in scholen.

In dit artikel volgt eerst een korte typering van de inhoud van de vernieuwing basisvorming als een grootschalige onderwijsverandering in het Nederlandse onderwijs (paragraaf 2). Daarna geven we een beknopte innovatietheoretische uiteenzetting die dient als kader voor de uitvoering van het onderzoeksproject. Paragraaf 3 bevat een korte verantwoording van de opzet van het onderzoek - aangeduid als Bavo-project - en de gehanteerde methoden. De onderzoeksvragen en de theoretische verwachtingen worden daarbij in kaart gebracht. In paragraaf 4 geven we een beeld van de beginsituatie en de activiteiten in het eerste voorbereidingsjaar van de drie scholen, die in het onderzoek centraal staan. Dit artikel bevat het verslag van de eerste onderzoeksperiode van 1991 tot eind 1992.

2 De basisvorming als innovatie

2.1 De inhoud en de strategie

De basisvorming houdt *inhoudelijk* gezien een complexe verandering in van het onderwijs in de eerste leerjaren van het voortgezet onderwijs. De hoofdbestanddelen van de innovatie inhoud zijn:

- Elke school moet een vast pakket van 15 vakken aanbieden (Nederlands, Engels, Frans of Duits, geschiedenis en staatsinrichting, aardrijkskunde, economie, wiskunde, natuur- en scheikunde, biologie, informatiekunde, techniek, lichamelijke oefening, verzorging, keuze uit: muziek/dans/drama, keuze uit: tekenen/handvaardigheid/textiele werkvormen.
- Een adviestabel is opgesteld voor het aantal uren dat per vak gegeven *kan* worden. Scholen zijn niet aan deze tabel gehouden.
- Het aantal lesuren wordt uitgebreid van 30 naar 32 uur per week.
- De scholen mogen 20% van het totaal aantal lesuren vrij besteden.
- Er zijn kerndoelen geformuleerd per vak. Deze kerndoelen moeten worden getoetst aan de hand van toetsen die de overheid beschikbaar stelt. Scholen mogen in uitzonderingsgevallen ook vervangende toetsen maken. Elke school bepaalt zelf wanneer de toetsen worden afgenomen, ergens tussen het einde van het tweede en vierde leerjaar.
- Voor de kerndoelen per vak en de vakoverstijgende doelen en vaardigheden geldt dat zij gekenmerkt zijn door 'toepassing, vaardigheden en samenhang'.

De *strategische* aspecten van deze vernieuwing zijn voor een deel door de overheid bepaald. Nadat het wettelijk kader is vastgesteld krijgen scholen de opdracht in een tijdsbestek van twee jaren de invoering in de onderwijspraktijk voor te bereiden. De basisvorming is te beschouwen als een poging om vanuit de centrale overheid door middel van nieuwe wetgeving en ondersteunende maatregelen een aantal algemene doelstellingen voor de eerste fase van het voortgezet onderwijs te verwezenlijken. We typeren dit als een *onderwijsinnovatie*, gedefinieerd als *een planmatige poging de kwaliteit van het onderwijs in een of meer scholen te verbeteren op grond van vastgestelde doelen* (vgl. Lagerweij, 1987, p. 110). De voorgestelde basisvorming is een grootschalige innovatie, die voor elke school in het voortgezet onderwijs als geheel en voor alle leraren daarbinnen werkzaam tot wijzigingen kan leiden (over grootschaligheid zie Van den Berg & Vandenberghe, 1984).

In augustus 1991 is het invoeringsproces van basisvorming voor alle scholen voor voortgezet onderwijs gestart. Voordat de implementatie een aanvang zal nemen, dient elke school na te gaan op welke wijze men de door de wet vereiste veranderingen zal gaan concretiseren. Het betreft een reeks van beslissingen op het gebied van onderwijsinhouden (vakkenpakket, vakkenintegratie) en schoolorganisatie (o.a. lesrooster, opzetten leerlingbegeleiding).

2.2 Het theoretisch kader

Het totaal aan theorieontwikkeling dat sinds de jaren zestig inzake onderwijsinnovatie werd opgebouwd, is niet van dien aard dat op grond daarvan van te voren met grote mate van zeker-

heid valt aan te geven hoe veranderingen in scholen zich zullen gaan voltrekken en op welke effecten men na verloop van tijd kan rekenen (zie voor recente overzichten Lagerweij & Voogt, 1990; Hopkins, 1990; Fullan, 1991; Louis & Miles, 1990; Dalin, 1992). De voorspellende kracht van het innovatietheoretisch kennisbestand is beperkt. Dat is een gevolg van de complexiteit van innovatieprocessen in het onderwijs. Het blijkt steeds weer, dat zeer veel factoren het verloop van veranderingsprocessen beïnvloeden. Deze factoren hangen onderling vaak sterk samen. Dat verschijnsel leidt er ook toe dat veranderingen in het onderwijs moeilijk van buitenaf te sturen zijn (zie tevens Fullan, 1991, p. XII).

Het onderzoek waarover in de innovatieliteratuur gerapporteerd wordt, is doorgaans gericht op analyse en interpretatie achteraf van het daadwerkelijk verloop van veranderingsprocessen in het onderwijs in scholen. Dit onderzoek heeft geleid tot meer zicht op welke factoren in scholen nu werkelijk invloed hebben op het realiseren van gewenste veranderingen. Tegelijkertijd is duidelijk geworden dat zich hierbij verschijnselen voordoen, die in elke context een eigen verloop laten zien. "The number and dynamics of factors that interact and affect the process of educational change are too overwhelming to compute in anything resembling a fully determined way. (..) there are no hard-and-fast rules, rather a set of suggestions or implications given the contingencies specific to local solutions" (Fullan, 1991, p. 47).

In het Bavo-project gaan de onderzoekers uit van de volgende innovatietheoretische uitgangspunten:

1. Invoering van basisvorming is op te vatten als een proces van schoolontwikkeling. *Schoolontwikkeling is het op planmatige wijze voortdurend sturing geven aan de inrichting van het onderwijs in de school en aan interne voorwaarden teneinde de vastgestelde doelen te bereiken.* (Vgl. voor het begrip schoolontwikkeling ook Van der Krogt, 1983; Hofman & Haanstra, 1986; Voogt, 1989; Lagerweij, 1993b). Hieruit volgt dat onderzoek naar het invoeren van basisvorming gericht dient te zijn op de school als geheel èn op de betrokkenen afzonderlijk als eenheid van verandering.

2. Voor het welslagen van de basisvorming is het van belang dat scholen *planmatig* te werk gaan op basis van de beoogde doelstellingen. Betrokkenen dienen gelegenheid te krijgen tot leren, dat wil zeggen het opdoen van de benodigde kennis en het verwerven van noodzakelijke vaardigheden. Daarvoor zijn naast druk (van overheid, bestuur, schoolleiding, collega's) ook vormen van ondersteuning (nascholing, begeleiding, leermiddelen) gewenst en noodzakelijk.

3. Onderwijsverandering is een *evolutionair* proces. De veranderingen voltrekken zich in een tijdsbestek waarvoor termijnen van meer dan vijf jaar normaal zijn (Dalin, 1992, spreekt zelfs van tenminste tien jaar). Het is mogelijk en gebruikelijk in dat proces een onderscheid te maken tussen een drietal subprocessen, te weten adoptie, implementatie en continuering. Diverse factoren kunnen de voortgang van het proces belemmeren of bevorderen (zie voor een overzicht Fullan, 1991).

4. Er is bij de invoering van basisvorming een verband tussen de veranderingen in de microsituatie (klasniveau) en de *schoolorganisatie*. De onderwijskundige vormgeving van basisvorming stelt eisen aan het organisatorische subsysteem (schoolleiding, management, vaksecties e.d.) en aan de betrokkenheid van de leraren (Marx, 1975; Van den Berg & Vandenberghe, 1981; Fullan, 1992).

5. Voor de realisering van basisvorming moet de school een beroep doen op *veranderingscapaciteiten*. Dit begrip verwijst naar vermogens van de school (dat wil zeggen de schoolleiding en het schoolteam) om veranderingen te kunnen plannen, uitvoeren en begeleiden met het oog op onderwijsinhoudelijke ontwikkeling en de daarvoor noodzakelijke ontwikkeling van de schoolorganisatie (vergelijk Hofman & Haanstra, 1986; Van Gennip, 1991).

6. Bij de invoering van basisvorming is het te verwachten dat per school een *eigen vormgeving* ontstaat. Lokale situatie (ouders, bestuur, omgeving), schoolkenmerken (grootte, type, gebouw), de schoolcultuur (zie Staessens, 1990) en het schoolconcept (Lagerweij, 1993a) beïnvloeden de implementatie en de leerresul-

taten van basisvorming. Naast het in kaart brengen van deze implementatie is het eveneens gewenst de leereffecten en niet voorziene effecten te meten.

Op grond van deze uitgangspunten is voor het onderhavige project een aantal verwachtingen opgesteld, die in paragraaf 3.2 vermeld worden. Daar is tevens het conceptuele kader van het project weergegeven.

2.3 Voorafgaand Nederlands onderzoek

Hoe zullen scholen erin slagen tijdig en verantwoord in 1993 met de implementatie te starten? Hoe zal basisvorming er in praktijk gaan uitzien? Welke effecten zullen na verloop van tijd zichtbaar zijn?

Op dergelijke vragen is op grond van het tot nu toe in Nederland verrichte innovatieonderzoek vooraf moeilijk een steekhoudende voorspelling te doen. Sinds de jaren zeventig zijn in ons land ervaringen opgedaan met grootschalige projecten die enigszins vergelijkbaar zijn met de invoering van de basisvorming: het Mavo-project, de Middenschoolexperimenten en het Integratieproject Scholengemeenschappen. Deze projecten zijn in 1985 opgegaan in een overkoepelend project, het project Scholen in Ontwikkeling (SiO-project)[2].

In 1987 is door Hofman, De Canne, Olthof en Bos (1988a) een beleidsvoorbereidende studie uitgevoerd naar de rol van het onderwijsmanagement bij de implementatie van een grootschalige onderwijsvernieuwing. Zij zijn nagegaan welke ervaringen uit voorgaande projecten voor de basisvorming conclusies opleveren. In deze studie is een aantal deskundigen geraadpleegd, zijn grootschalige projecten geanalyseerd en heeft men buitenlandse ontwikkelingen in de eerste fase van het voortgezet onderwijs bestudeerd. Voorts wordt in een tweetal publikaties (Hofman e.a., 1988b, 1988c) verslag gedaan van case-studies in smalle en brede scholengemeenschappen.

Uit de case-studies (zie Hofman e.a., 1988b) blijkt onder meer dat de vernieuwingen stagneerden op die scholen, waar het organisatorisch systeem onvoldoende ontwikkeld is in vergelijking met de beoogde vormgeving van het onderwijskundig systeem. De onderzoekers wijzen er verder op dat het veranderen van het onderwijs in de scholen afhangt van de beschikbaarheid van geschikte leermiddelen, mede in relatie tot heterogeen of homogeen groeperen. Ook moeten leermiddelen de scholen helpen vorm te geven aan de mogelijke verschillen in tijdsduur (2 à 4 jaar) die leerlingen ter beschikking hebben om de doelen van de basisvorming te bereiken. Voorts geven zij het belang aan van een systeem van leerlingbegeleiding. Uit de eerdergenoemde projecten menen de onderzoekers te kunnen vaststellen, dat fusieprocessen een vertragende invloed hebben op het verloop van andere vernieuwingen in scholen. Ten slotte is, aldus de onderzoekers, gebleken dat het wenselijk is dat scholen de beschikking hebben over interne en externe begeleiding. Deze begeleiders moeten hun activiteiten (nascholing) afstemmen op de behoeften van de scholen.

Voor het invoeren van basisvorming zijn niet in alle scholen de genoemde voorwaarden even gunstig. Hofman en Lugthart (1991) hebben bij een inventariserend onderzoek in 243 scholen getracht vast te stellen welke interne capaciteiten in scholen voor voortgezet onderwijs aanwezig zijn met het oog op de invoering van basisvorming. De onderzoekers achten deze steekproef "een redelijke afspiegeling van de landelijke populatie" (o.c., pag. 78). Het bleek dat de huidige situatie nog ver verwijderd is van het streefbeeld van basisvorming in brede schoolgemeenschappen. "Bij een kwart tot een derde deel van de scholen is sprake van een planmatig innovatiebeleid. Bij die scholen is sprake van:

- een opgesteld document waarin de vernieuwingsplannen van de school op de langere termijn zijn beschreven;
- een taakverdeling voor invoering van de innovatie;
- een tijdsplanning voor invoering van de vernieuwing;
- een opgesteld invoeringsplan;
- een specifiek orgaan dat als 'motor' voor vernieuwingen fungeert;
- een beleid waarbij alle geledingen in de school regelmatig betrokken worden bij innovaties (66%)." (o.c., pag. 106).

Interne capaciteiten die volgens de auteurs een positieve bijdrage leveren aan schoolontwikkeling in de richting van basisvorming zijn: een gedegen coördinatie in de school door een geïntegreerd professionaliseringsbeleid, een

planmatig innovatiebeleid èn de ontwikkeling van een schoolcultuur waarin sprake is van een eenduidige regelgeving en een optimaal op onderwijs en prestatie gerichte leeromgeving.

Er is dus zeker reden aan te nemen dat voor het gehele scholenveld de uitgangssituatie voor invoering van de basisvorming weinig gunstig is. Hofman en Lugthart (1991) stellen vast dat brede scholengemeenschappen voorlopers zijn op alle drie ontwikkelingsgebieden (i.c. curriculum, differentiatie en leerlingbegeleiding) van de basisvorming. De categoriale scholen hebben de langste weg te gaan. Smalle scholengemeenschappen zitten hier tussen in. Slechts een groep van 8% van de scholen zal met geringe extra inspanning de veranderingen kunnen invoeren. Een derde deel van de scholen staat nog ver af van realisering van de gewenste vernieuwingen.

In bovengenoemd onderzoek zien we overeenkomstige innovatietheoretische noties als in paragraaf 2.2 door ons zijn opgesteld (bijvoorbeeld schoolontwikkeling, contingentiebenadering, veranderingscapaciteiten, het proceskarakter van innovaties, noodzaak tot afstemming onderwijskundige en organisatorische subsystemen, het belang van leermiddelen).

Er zijn ook enkele verschillen in de opzet en inhoud van het onderzoek. De genoemde onderzoekers konden tijdens hun werkzaamheden niet volledig de laatste besluiten betreffende de basisvorming verwerken. Sinds 1985 zijn er verschillende wetsvoorstellen geweest, die uiteenlopende uitwerkingen bevatten van o.m. de duur van de basisvorming, het aantal niveaus, de eindtermen of kerndoelen, het aantal vakken en het invoeringstraject. Daardoor zijn enkele onderzoeksgegevens inmiddels niet meer relevant. Het Utrechtse Bavo-project sluit aan bij de definitieve wettelijke regelingen.

De RION onderzoekers hebben geen longitudinaal design gehanteerd. De processen van verandering zijn niet in de tijd gevolgd. Er is volstaan met eenmalige metingen met behulp van interviews. Er is bovendien geen onderzoek verricht op klasniveau.

3 Het onderzoeksdesign

3.1 Probleemstelling

Het voorgaande bevat voldoende redenen voor het opzetten van een longitudinaal onderzoek naar de wijze waarop in scholen het concept basisvorming vorm krijgt. Daartoe is voor het Bavo-project de volgende algemene vraagstelling geformuleerd:

In welke mate en op welke wijze dragen factoren op meso- en microniveau bij aan de invoering en het bereiken van de doelstellingen van de basisvorming in scholen voor voortgezet onderwijs en hoe is deze bijdrage te verklaren?

Het betreft een beschrijvend en interpreterend onderzoek naar condities, processen en effecten van een door de overheid geïnitieerd innovatieproject waaraan alle scholen in het voortgezet onderwijs moeten deelnemen. De condities buiten de scholen - onderwijsbeleid, ondersteuning e.d. - worden in het onderzoek betrokken voor zover deze van invloed zijn op het verloop van de veranderingsprocessen.

Als uitgangspunt voor dit onderzoek geldt dat invoering van de basisvorming leidt tot veranderingen in de school als organisatie en bij de vormgeving van onderwijsleerprocessen door de leerkrachten. Dat betekent dat er sprake moet zijn van een multi-level aanpak en analyse.

De algemene vraagstelling is onderverdeeld in deelvragen, die geordend zijn in 3 clusters. Hierna geven we globaal in trefwoorden aan welke aspecten nader onderzocht worden.

(1) De schoolkenmerken

De beginsituatie van de school (schoolcultuur, samenstelling docentenkorps, leerlingenpopulatie, veranderingscapaciteiten, doorstroom- en uitstroomgegevens, de lokale omgeving, de organisatie/structuur, het management, opvattingen over basisvorming).

(2) Het procesverloop (incl. ondersteuning en evaluatie)

Het verloop van de besluitvorming tijdens de adoptie en de implementatie betreffende de vormgeving en de organisatorische en onderwijskundige problemen die daarbij op de verschillende niveaus in de schoolorganisatie

voorkomen (incl. vaksecties en het onderwijs in de klassen), de behoeften aan ondersteuning.

(3) Onderwijsleerproces en leereffecten

Het verloop van onderwijsleerprocessen in de klassen, klassesamenstelling (omgaan met heterogeniteit), de realisatie van kerndoelen en vakoverstijgende doelen (met name toepassingsgerichtheid en samenwerking), de veranderingen in leerinhouden, het gebruik van leermiddelen. Resultaten van de invoering van basisvorming in termen van leerprocessen en -resultaten (cognitief en affectief) bij de leerlingen, doorstroom en uitstroom van de leerlingen.

3.2 Conceptueel kader

In relatie tot de invoering van de basisvorming zijn verwachtingen opgesteld. Wij koesteren deze op grond van eerder onderzoek en theorievorming omtrent de ontwikkeling van basisvorming in de scholen. We rangschikken de verwachtingen aan de hand van de bovengenoemde drie clusters van variabelen.

(1) De schoolkenmerken

1.1. Het welslagen van de implementatie van basisvorming is voor een belangrijk deel afhankelijk van de mate waarin de visie van het personeel overeenstemt met de idealen van deze vernieuwing (Louis & Miles,1990; McLaughlin, 1990; Stokking, 1992).

1.2. De mate van individualisering die een school bij implementatie van basisvorming kan realiseren is afhankelijk van de mate waarin de organisatie horizontaal is (Van der Krogt, 1983; Giesbers & Van der Krogt, 1990).

1.3. De veranderingscapaciteiten bepalen in hoeverre scholen erin slagen met inzet van eigen management en eventuele externe ondersteuning de beoogde wijzigingen vorm te geven (Van Gennip, 1991).

1.4 Invoering van de basisvorming gaat sneller naarmate de school beschikt over coördinatiemechanismen, zoals geïntegreerd professionaliseringsbeleid, formalisatie van regelgeving (Hofman & Lugthart, 1991).

(2) Het procesverloop

2.1 Scholen zullen op uiteenlopende wijze invulling geven aan de implementatie van basisvorming al naar gelang hun actuele situatie (Hofman & Lugthart, 1991).

2.2 De veranderingscapaciteiten zijn rechtstreeks van invloed op de mate van implementatie van het vernieuwingsconcept basisvorming (Van Gennip, 1991).

2.3 Participatie van docenten in de besluitvorming werkt bevorderend zowel op de vormgeving van het onderwijs als op de effecten op leerlingniveau (Anderson, 1982; Rosenholtz, 1985; Oakes, 1987)

2.4 Een planmatig management (met instelling van werkgroepen en commissies met onderlinge afstemming, een beleidsplan, invoeringsplan) bevordert schoolontwikkeling en implementatie van de vernieuwing (Hofman & Lugthart, 1991, p. 48).

2.5 De invoering van basisvorming verloopt des te beter naarmate het onderwijzend personeel bereid en in staat is tegemoet te komen aan verschillen tussen leerlingen.

(3) Onderwijsleerproces en leereffecten

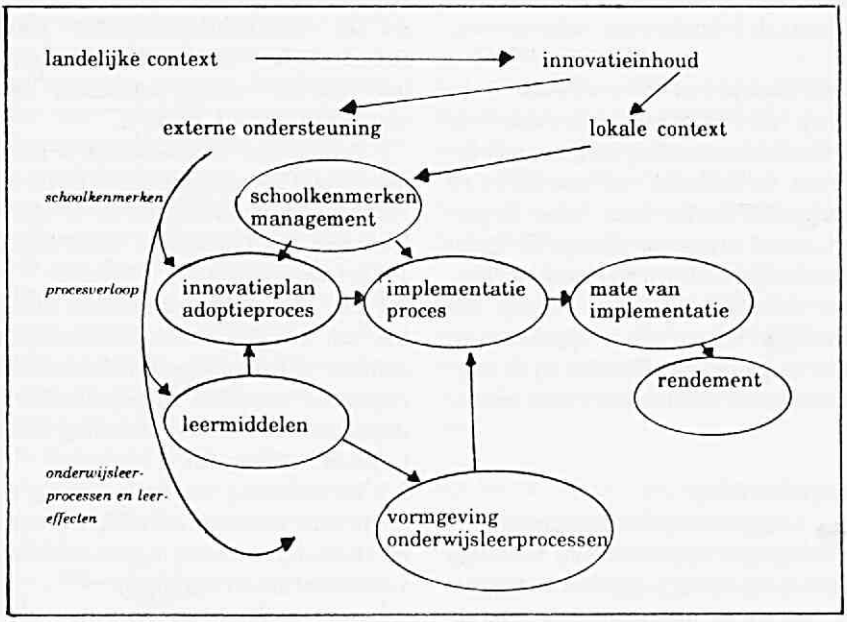
3.1 Invoering van basisvorming leidt tot wijziging van vormgeving van onderwijsleerprocessen bij alle schoolvakken.

3.2 Verhoging van het peil van jeugdonderwijs als gevolg van invoering van basisvorming treedt pas op als docenten gelegenheid gehad hebben enkele jaren met deze veranderingen te werken.

3.3 Extern ontwikkelde leermiddelen en externe begeleiding kunnen positieve invloed hebben op de implementatie en incorporatie van basisvorming (McLaughlin, 1990).

3.4 Invoering van de basisvorming leidt tot veranderingen in de mate waarin betrokkenen (docenten en leerlingen) zich verbonden voelen met de school als geheel en de klas(sen) als onderdeel daarvan.

De in de vorige paragraaf onderscheiden clusters van vraagstellingen zijn met het oog op de uitvoering van een onderzoek in Figuur 1 bijeengezet. Het betreft een afbeelding waarin tot uiting komt welke begrippen in het onderzoek betrokken zullen worden en welke relaties tussen de begrippen verondersteld worden. De overwegingen die ten grondslag liggen aan het model zijn reeds in het innovatietheoretisch kader (zie paragraaf 2) vermeld.



Figuur 1. Algemeen onderzoeksmodel van het Bavo-project

Door Miles en Huberman is een dergelijk kader aangeduid als 'conceptual framework'. Zij omschrijven dat als volgt: "A conceptual framework explains, either graphically or in narrative form, the main dimensions to be studied - the key factors, or variables - and the presumed relations among them." (Miles & Huberman, 1984, p. 28).

Figuur 1 geeft op grafische wijze aan welke drie clusters van variabelen in het onderzoek centraal staan. De pijlen geven aan op welke wijze zij met elkaar in verband staan. Dat betekent in concreto dat we ervan uitgaan dat de landelijke en lokale context de innovatieinhoud van de basisvorming opleveren. De externe ondersteuning verricht ook in dat kader dienstverlenende activiteiten ten behoeve van de scholen. Daarmee is de achtergrond van het vernieuwingsproces voor de afzonderlijke scholen gegeven.

Vervolgens gaan wij ervan uit dat het verloop van de veranderingsprocessen op schoolniveau nauw samenhangt met kenmerken van de schoolorganisatie (met name de veranderingscapaciteiten, de schoolcultuur en het management). Zowel tijdens de adoptie als tijdens de implementatie werken dergelijke kenmerken door en bepalen mede de in de loop van het proces verkregen effecten.

De bedoeling van de basisvorming is vooral via de kerndoelen de onderwijsleersituaties te veranderen. De basisvorming krijgt vorm in de onderwijsleerprocessen die de (vak)docenten in interactie met leerlingen en leermiddelen realiseren. In Figuur 1 is aangegeven dat de implementatie van deze vernieuwingen wordt gemeten aan de hand van wat zich in onderwijsleersituaties afspeelt. Tevens geven wij aan dat daarbij de kenmerken van de school als organisatie (bijvoorbeeld het overleg in vaksecties, collegiale samenwerking, optreden schoolleider, de wijze van besluitvorming) een rol spelen.

3.3 Onderzoeksopzet

In het onderzoek is vooral gekozen voor gevalstudies. Daarbij wordt gebruik gemaakt van onderzoektechnieken als documentanalyse, bijwonen van vergaderingen en bijeenkomsten, observaties in klassen, afnemen van toetsen, vragenlijsten en interviews. Op drie scholen zal gedurende vijf jaar worden nagegaan hoe de invoering van basisvorming verloopt. Het betreft drie brede scholengemeenschap-

pen, die zich bereid verklaard hebben aan het onderzoek in principe voor langere tijd mee te werken. De keuze voor brede scholengemeenschappen is ingegeven vanwege het feit dat zij als voorlopers kunnen worden beschouwd van scholen op weg naar basisvorming gelet op de brede instroom van leerlingen en een doorgaans breed vormingsaanbod.

Bij de bestudering van innovatieprocessen wordt vaak gebruik gemaakt van gevalstudies. Door het van dichtbij volgen van veranderingsprocessen in de dagelijkse werkelijkheid van scholen is men er allengs in geslaagd een rijk geschakeerd beeld van de complexiteit van onderwijsverandering te geven[3].

De argumenten die tot de keuze van deze onderzoeksbenadering voor het proces van invoering van basisvorming hebben geleid zijn:

- Basisvorming als vernieuwing betekent veranderingsprocessen in scholen met een complex en multidimensionaal karakter, dat wil zeggen dat op alle in de school te onderscheiden niveaus (inclusief personen) er wijzigingen kunnen en moeten optreden. Dat geheel moet in kaart worden gebracht om volledig zicht te krijgen op de implementatie (zie ook Beem, 1988).
- Basisvorming betekent een zich op langere termijn ontwikkelende reeks veranderingen in scholen. Het proces vraagt naar verwachting tenminste enkele jaren alvorens er duidelijk zicht is op de uitvoering in de onderwijspraktijk.
- De vormgeving van basisvorming kan verschillen van school tot school. De schoolspecifieke factoren en hun invloed moeten in beeld komen, ook in een wat langer tijdsbestek.
- De weinig operationeel uitgewerkte inhoud van de basisvorming laat vele interpretaties toe. De nadere invulling en de daarbij gevolgde besluitvorming moeten aan het licht komen.

Voor het vormgeven van een gevalstudie kan men vele wegen inslaan. Centrale opgave voor de onderzoekers is de ontwikkelingen in de tijd zo goed mogelijk te volgen waarbij de context waarbinnen de verschijnselen zich afspelen eveneens in het onderzoek betrokken wordt. Door Yin (1989, 1991) zijn belangwekkende bijdragen geleverd aan de techniek van het opzetten van een gevalstudie. Het onderhavige project is daarop gebaseerd. Yin wijst erop, dat het theoretisch belang van gevalstudies toeneemt, als de onderzoekers vooraf theoretisch onderbouwde verwachtingen formuleren In paragraaf 3.2 zijn de innovatietheoretische verwachtingen aangegeven die in het Bavo-project bestaan over de invoering van basisvorming.

In deze gevalstudie maken we gebruik van kwalitatieve en kwantitatieve technieken van dataverzameling en -verwerking. Deze combinatie is niet zo gebruikelijk in gevalstudies, maar levert naar onze opvatting meer informatie op.

Over het algemeen zijn, aldus Yin (1989), 'explanatory' case studies te verkiezen als er sprake is van 'hoe' en 'waarom' vragen, als de onderzoeker weinig controle heeft over de gebeurtenissen en als de gerichtheid is op een hedendaags verschijnsel dat zich afspeelt in een levensechte situatie. Hij definieert de case study als volgt:

"A case study is an empirical inquiry that:
* investigates a contemporary phenomenon within its real-life context; when
* the boundaries between phenomenon and context are not clearly evident; and in which
* multiple sources of evidence are used." (o.c. p. 23).

Ter aanvulling en verificatie van de gegevens uit de gevalstudie is het van belang dat herhaalde surveys gehouden worden onder een representatieve steekproef scholen die bij de basisvorming betrokken zijn. Begin 1993 heeft het Proces Management Basisvorming daarover een eerste rapportage gegeven (zie Voogt, 1993).

4 De resultaten van de gevalstudies

Het onderzoek heeft plaats in drie brede scholengemeenschappen van verschillende signatuur (openbaar en twee bijzonder onderwijs). Voor wat betreft de schoolkenmerken zijn basisgegevens aan de drie scholen gevraagd. Tevens is gebruik gemaakt van interviews en vragenlijsten. Geïnterviewd zijn enkele ouders, leerlingen, docenten in de vakken wiskunde, Engels en muziek en de schoolleiding. Met be-

hulp van het SAS instrumentarium (Systematische Analyse voor Schoolontwikkeling, zie Voogt, 1989) zijn verschillende schoolkenmerken in kaart gebracht. De vragenlijsten voor SAS zijn bestemd voor het gehele personeel. De leerlingen hebben vragenlijsten ingevuld over samenhang (hechtheid) in de klas en de houding ten opzichte van de school (schoolgerichtheid).

Over elke school is een vertrouwelijk rapport opgesteld, dat ter fiattering aan de schoolleiding is voorgelegd. Gedurende het onderzoek zal daarover een samenvatting gepubliceerd worden (zie ook paragraaf 4.1).

Het procesverloop in de adoptiefase is in kaart gebracht met behulp van documenten die daarop per school betrekking hebben. Bovendien is door regelmatig bezoek van onderzoekers aan de scholen en door het bijwonen van speciale vergaderingen over basisvorming een zo getrouw mogelijk beeld van de ontwikkelingen per school verkregen.

Over het derde cluster variabelen, de inrichting van onderwijsleerprocessen, is in deze onderzoeksfase nog nauwelijks informatie verzameld. Vlak voor de start van het schooljaar 1993-1994 zijn daartoe beginmetingen gepland die betrekking hebben op onder meer de plannen, oordelen en problemen bij het personeel rondom de eigen uitwerking van basisvorming. Daarna starten onderzoeksactiviteiten die gericht zijn op de uitvoering van basisvorming op klasniveau.

In Figuur 2 zijn de onderzoeksthema's voor zover tot nu toe onderzocht met de bijbehorende instrumenten bijeen gebracht. De volgende paragrafen bevatten een overzicht van de verkregen gegevens.

4.1 Schoolkenmerken

We geven een korte karakteristiek van de scholen, aangeduid als school A, B en C.

School A

Het betreft een openbare scholengemeenschap met een IVBO-afdeling, VBO, MAVO, HAVO en Atheneum. De school opende in 1983 haar poorten. De school staat in een sociaal arme buurt waarin zich veel problemen voordoen. Er zijn relatief veel probleemleerlingen; 9% is allochtoon. De school is een concentratieschool voor anderstaligen. De IVBO-afdeling betreft kinderen met leermoeilijkheden (mlk), geen gedragsmoeilijkheden (mok).

De school groeit nog steeds. Het schoolgebouw is te krap en er moet goed op vandalisme

cluster variabelen	variabelen	instrumenten	*respondenten*
schoolkenmerken	schoolleiding schoolorganisatie schoolklimaat onderwijsaanbod visie prestatiegerichtheid randvoorwaarden en middelen samenstelling personeel leerlingen instroom en doorstroom gegevens	SAS vragenlijst interviews administratieve gegevens	alle docenten onderwijsondersteunend personeel schoolleiding
procesverloop	takenpakket wijze van invoering basisvorming planning opinies over basisvorming oordelen over de school	interviews documenten analyse	schoolleiding coördinatoren docenten onderwijsondersteunend personeel ouders leerlingen
onderwijsleerprocessen	schoolgerichtheid hechtheid	vragenlijsten	leerlingen brugjaar

Figuur 2. Clusters variabelen met bijbehorende instrumenten voor beginmeting

worden gelet. In 1992 is een nieuwe rector aangetreden.

School A heeft een breed vormingsaanbod. Het onderwijs is leerlinggericht en er is sprake van structurele leerlingbegeleiding. Heterogeniteit staat hoog in het vaandel. De school wordt democratisch geleid en er zijn veel mogelijkheden voor initiatief. De school is betrokken geweest bij het project Scholen in Ontwikkeling (SiO).

Er zijn 112 docenten, 5 directieleden en 20 oop-leden. Er bestaat een tweejarige heterogene brugperiode. Daarna verdeelt men de leerlingen in een stroom voor VBO/MAVO en een stroom voor HAVO/VWO. Na klas 3 splitst men het HAVO/VWO-deel in twee stromen: één voor HAVO en één voor VWO.

School B

School B is een brede rooms-katholieke scholengemeenschap met IHNO, LHNO, MAVO, HAVO en Atheneum. De school is in 1985 ontstaan uit een fusie van een categoriale LHNO en een categoriale MAVO (later uitgebreid met HAVO en VWO).

Het fusieproces is in de scholengemeenschap nog voelbaar. Er is sprake van het bijeenkomen van verschillende schoolculturen. Er is een soort tweedeling in het docentenkorps volgens de samenstellende categoriale scholen uit het dorp en de nabijgelegen stad.

De leerlingpopulatie komt voor 85% uit de plaats, de overige 15 % komt uit twee naburige plaatsen. Pogingen om meer leerlingen uit de regio aan te trekken stuiten op denominatie- en cultuurverschillen. Tweederde van de leerlingpopulatie bestaat uit meisjes, éénderde uit jongens. Dit komt doordat de school een LHNO heeft en een IHNO. Er zijn geen allochtone leerlingen.

Het specifieke karakter van de school is toe te schrijven aan het karakter van de plaatselijke bevolking. Er is sprake van een hechte gemeenschap, een zeer sterke sociale controle en dientengevolge van groepsgedrag.

Na de fusie op 1-8-1985 waaruit school B ontstond heeft de school tot 1988 deelgenomen aan het project Scholen in Ontwikkeling (SiO).

In 1992 is een nieuwe rector aangetrokken. De school heeft een tijdlang zonder aanvullende faciliteiten met een onvolledige directie gefunctioneerd. Er is in 1991 sprake van een moeizame verhouding met het schoolbestuur. Bestuurswisseling heeft geleid tot betere verhoudingen.

De gebouwensituatie is goed. De school heeft 73 docenten, 6 directie- en 12 oop-leden. De organisatiestructuur kent brugklassen volgens een 'dakpansysteem' ingedeeld, nl. HAVO/Atheneum, MAVO/HAVO, LBO/MAVO en een aparte IHNO-klas. In het tweede jaar gaat men streamen. HAVO/Atheneum blijft gemixed tot en met het derde jaar bij elkaar.

School C

Het betreft een scholengemeenschap VBO/VWO op protestant-christelijke grondslag. De identiteit is weloverwogen gekozen, speelt in alle facetten van het schoolleven een beslissende rol, is door alle betrokkenen geaccepteerd en is voor de meeste ouders de voornaamste reden waarom voor deze school is gekozen. Ook het benoemingsbeleid is hierop gebaseerd. Hoewel de school pas betrekkelijk korte tijd bestaat (1984) is het aantal leerlingen toch al rond de 1200. Leerlingen komen uit een wijde omtrek van de school. De school groeit nog jaarlijks, al is over enige tijd wel een stabilisatie te verwachten. Het aantal allochtone leerlingen is beperkt (hooguit 10 per jaar). Het gaat dan meestal om adoptiekinderen.

Het schoolgebouw is nog maar kort in gebruik. Voordien bestond de school uit verschillende lokaties. De klassen zijn homogeen samengesteld. De meeste leraren geven klassikaal onderwijs. De selectie van de leerlingen gebeurt vooral op grond van het advies van het hoofd van de basisschool. Met Kerstmis en Pasen wordt de indeling gecorrigeerd, waarbij respectievelijk 8% en 4% wordt herplaatst. Bij de organisatie van het lesgeven is de groep uitgangspunt, niet de individuele leerling. De leerstofplanning ligt vast. Binnen klasseverband wordt niet gedifferentieerd.

De directie bestaat uit 4 leden en neemt een centrale plaats in. Er zijn 80 docenten werkzaam en er zijn 17 oop-leden. De secties functioneren relatief autonoom. Er is sprake van een open, positieve sfeer en er zijn veel informele contacten onder het personeel. De individuele leerlingbegeleiding is beperkt.

De kerndoelen worden, in overleg met andere scholen op dezelfde grondslag, tegen deze achtergrond beoordeeld en ingevuld. In het eerste brugjaar worden leerlingen dakpansgewijze ingedeeld. De keuze voor homogeen onderwijs is ook levensbeschouwelijk gefundeerd. De leerlingen genieten veel vertrouwen. Ten aanzien van basale inzichten en opvattingen bestaan er tussen leraren geen verschillen van mening.

SAS beelden

De bedoeling van het SAS-instrument (zie Voogt, 1989) is het in kaart brengen van de totale school. Alle betrokken schoolleiders en docenten geven aan hoe zij over diverse aspecten van het functioneren van de school denken. Tevens wordt gevraagd aan te geven in hoeverre men tevreden is met de bestaande situatie en of actie voor het veranderen van bepaalde zaken wenselijk is. In Figuur 3 is aangegeven hoe de drie scholen scoren op 7 schalen die in het instrumentarium zijn opgenomen, te weten

Schoolleider; deze factor is de weergave van de mate waarin de docenten het functioneren van de schoolleider als stimulerend ervaren; hierbij zijn betrokken zaken zoals het 'feed-back' geven aan docenten, het stimuleren van docenten tot eigen initiatieven en nascholing, het initiëren van vernieuwingen, het coördineren van de organisatie en het gemakkelijk benaderbaar zijn van de schoolleider.

Organisatie; dit geeft een beeld van de mate waarin de school in staat is zichzelf goed te organiseren, zoals goede coördinatie en afstemming van commissies, raden en functies, duidelijkheid in taken, bevoegdheden en regels, adequate doorstroom van informatie, goede vergaderfrequentie, gezamenlijke besluitvorming en een adequate opstelling van het lesrooster.

Sfeer; deze factor heeft betrekking op zaken zoals het kunnen uiten van ongenoegen, op de hoogte zijn van elkaars mening, het afwezig zijn van conflicten, gevoelens van onrust en onzekerheid, het geven van 'feedback' aan elkaar, sfeer tussen leerlingen en docenten, inspraak van leerlingen, beperkt verzuim van docenten en rivaliteit tussen secties.

Onderwijs; deze factor heeft betrekking op de flexibiliteit van het onderwijsaanbod van de school, zoals differentiatie, integratie van vakken, afstemming en aansluiting op behoeften en interesse van leerlingen en op maatschappelijke ontwikkelingen, het bespreken van het onderwijsaanbod, schoolwerkplanontwikkeling en extra-curriculaire zaken.

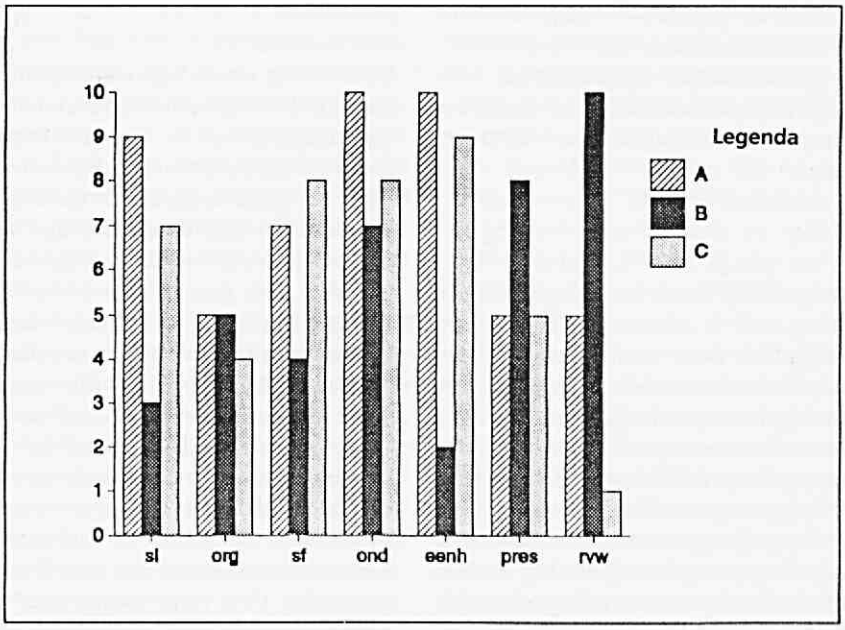
Eenheid; het gaat hier vooral om de gezamenlijke betrokkenheid op het onderwijs en de mate van eensgezindheid in de meningen over het onderwijs.

Prestatie; in welke mate creëert de school gunstige voorwaarden voor het behalen van hoge leerlingresultaten, zoals het bijhouden van leerlingresultaten, nadruk op goede prestaties, selectie-maatregelen en rapportage, consequent hanteren van toelatingseisen, tevredenheid over leerlingresultaten, het geven van huiswerk, de motivatie van de leerlingen, competentie van collega's, weinig ordeproblemen en goede afstemming op onderwijsdoelen.

Randvoorwaarden en middelen; het gaat hier om de tevredenheid over de materiële condities waaronder de school werkt.

In Figuur 3 zijn de scores van de drie scholen weergegeven. We gaan daarbij uit van decielscores. De decielen geven percentages scholen aan uit een landelijke steekproef die boven en onder het betreffende getal scoren. Zo betekent het zevende deciel dat 30 procent van de scholen een hogere score behaalt dan deze school en 70 procent lager uitkomt. Door het plaatsen van de drie schoolprofielen in de context van de landelijke situatie wordt de huidige stand van zaken duidelijk.

De drie projectscholen tonen op alle schalen onderling duidelijke verschillen. Aanzienlijke verschillen doen zich voor bij de schaal schoolleiding. School A en C scoren ruim boven het gemiddelde en school B zit twee decielen daaronder. De organisatie toont grote overeenkomsten tussen de scholen en alle drie zijn ze net aan het gemiddelde of een deciel daaronder (school C). De sfeer in A en C is ruim boven het gemiddelde en in school B een deciel daaronder. In de schaal onderwijs scoort school A uitstekend, terwijl de andere twee ook goed te noemen zijn. Bij de schaal eenheid zijn de onderlinge verschillen zeer groot. Vooral school B blijft hier ver achter bij de andere twee scholen en zit ook ver onder het gemiddelde. Bij de schaal prestatiegerichtheid scoort school B echter weer veel hoger dan de andere twee. Bij



Figuur 3. Decielscores van de 3 scholen op SAS factoren

de randvoorwaarden en middelen is school B uitstekend, terwijl school C zeer ver in negatieve zin afwijkt van het gemiddelde. School A zit hier juist op het gemiddelde.

Tenslotte blijkt dat alle drie scholen op één factor, namelijk onderwijs, boven het vijfde deciel uitkomen. A zit in het hoogste deciel, C in het achtste en B in het zevende deciel. In alle andere schalen blijft er tenminste één school staan op het vijfde of een daaronder liggend deciel.

4.2 Het procesverloop

We geven de stand van ontwikkeling van de adoptie van basisvorming weer aan het eind van 1992.

School A

De eerste periode (tot april) ging de informatievoorziening over basisvorming vooral schriftelijk (speciaal Bavo bulletin). Men vindt veel nog niet duidelijk, maar verwacht ook niet veel problemen, want 'we lopen voor' en men weet wel ondersteuning in te schakelen. Er bestaat het gevoel dat de basisvorming aansluit bij wat de school als concept hanteert. Enkele stemmen willen de vernieuwing gebruiken om ook extra aandacht aan de betere leerling te besteden.

De school heeft in de voorbereiding van de invoering Basisvorming de aanpak gekozen van het instellen van een projectgroep Basisvorming, die het voortouw neemt en plannen ontwerpt. Men is begonnen met de stand van zaken vast te stellen (SAS-rapport is daarbij leidraad). Daarna volgde een discussie omtrent de 'missie' van de school. De uitgangspunten c.q. het schoolconcept (gebaseerd op uitgangspunten voor de middenschool) is hier startpunt en niet bijvoorbeeld het invullen van de lessentabel (het rechtspositionele aspect). Duidelijk bleek dat deze discussie meerdere ronden vergt. In het schooljaar '91-'92 is deze discussie begonnen en in het volgende schooljaar is de school er mee doorgegaan. Er heerst een sfeer dat men tijd genoeg heeft om goed gefundeerd en goed voorbereid de basisvorming te kunnen invoeren per 1 augustus 1993.

Gunstige factoren voor een succesvolle invoering van de basisvorming zijn flexibiliteit van onderwijs, eenheid in onderwijs-ideeën en een stimulerend onderwijs-inhoudelijk leiderschap. Kwesties als heterogeen groeperen en differentiëren binnen klasseverband zijn voor deze school een kwestie van traditie (en ook breed gedragen: minimaal 80%) volgens de SAS-rapportage. Ook de leerlingbegeleiding, inherent aan de Basisvorming is voor deze

school een 'oud' gegeven. Sommige docenten vinden de Basisvorming 'een stap achteruit': minder (differentiatie) mogelijkheden voor i-leerlingen en ook voor VBO-leerlingen. Men acht een gevaar voor ongewenste 'streaming' aanwezig.

School B

Ter voorbereiding van de invoering van de basisvorming heeft de school een Bavo-projectgroep ingesteld. Deze overkoepelt acht subgroepen, die een bepaald thema bestrijken en van daaruit de consequenties van invoering bestuderen. Een ruime meerderheid van het totale docentenkorps is hierdoor direct betrokken bij de vormgeving van de basisvorming in de school. De projectgroep en de bijbehorende subgroepen doen regelmatig verslag van hun activiteiten. Daarbij heeft de beoogde vernieuwing meteen al aanleiding gegeven om tot explicitering van de schooldoelstellingen te komen. Daarbij is tevens zorg gedragen voor een systematische informatievoorziening, zowel over wat op de school zelf gebeurt als over inhoudelijke ontwikkelingen in en om de basisvorming zelf.

Informatiebulletins over de basisvorming zijn verspreid en de projectgroep basisvorming heeft een eindrapportage opgesteld (15.09.1992). De school heeft de voorbereiding systematisch aangepakt en taakuren beschikbaar gesteld voor de diverse projectgroepen. Voor diverse docenten is de basisvorming toch nog een zaak van de projectgroepen, al worden de consequenties soms met zorg tegemoet gezien. Het streven is om de basisvorming te gebruiken als een gelegenheid om de school verder te ontwikkelen.

Tegenover de visie van de schoolleiding van basisvorming als vernieuwingsproject staan de meer gereserveerde meningen van de leerkrachten. Naast erkenning van de algemene doelstellingen is er sprake van afwachten en scepsis. Men wacht af hoe de slag om de roosteruren zal uitvallen, hoe de methodenontwikkelaars hun werk zullen doen en wat er van het oorspronkelijke elan uiteindelijk zal overblijven. Verder is er scepsis ten aanzien van de pretentie van de vernieuwing dat alle kinderen zullen krijgen wat zij nodig hebben.

School C

De invoering van de basisvorming wordt voorbereid door zowel een interne als externe ondersteuningsgroep. De vraagstukken die met de invoering samenhangen zijn vooral voorwerp van gesprek binnen de directie. Omdat andere zaken (o.a. de nieuwbouw) veel aandacht vragen, is het moeilijk voldoende tijd vrij te maken. Eén docent heeft een coördinatietaak met betrekking tot de basisvorming. Een hoofdactiviteit betreft het verzamelen van informatie (o.a. via het tijdschrift van de landelijke coördinatiegroep) en doorgeven er van. Er is een intern mededelingenblad BAVO-info. Deelname aan de Interne projectgroep Bavo kost meer tijd dan aan taakuren is toegewezen. De docenten zijn slechts ten dele op de hoogte van de ontwikkelingen met betrekking tot basisvorming. Deze vernieuwing 'leeft' nog niet. Veel docenten hebben nog geen uitgesproken mening. Ook binnen de secties is de meningsvorming nog in een beginstadium.

Aan voorlichting is een en ander gedaan door middel van het oprichten van interne projectgroepen en een bezinningsdag. Er bestaat een overwegend positieve en welwillende houding ten aanzien van de basisvorming. De meningsvorming ten aanzien van vakinhoudelijke en vakdidactische zaken verkeert nog in een beginstadium; over de gevolgen van de invoering bestaat nog veel onzekerheid.

Samenvatting Bavo voorbereidingen

Vanuit innovatief gezichtspunt vertonen de scholen bij de adoptie van basisvorming naast verschillen ook overeenkomsten. Alle drie scholen hebben desbetreffende informatie op ruime schaal aangeschaft (handboeken, tijdschriften e.d.). In alle gevallen is ook een projectgroep basisvorming ingesteld, die uit meer dan 6 personen bestaat (school C heeft een kleinere groep). Daarnaast bestaan er in elke school werkgroepen van docenten die zich met aspecten van de basisvorming bezig houden. De vrije ruimte is in principe per school ingevuld, waarbij in alle scholen aandacht besteed wordt aan leerlingbegeleiding en studielessen (school B opteert ook voor klassieke talen, school A voor een derde vreemde taal). Naschollingsactiviteiten hebben eveneens op deze scholen plaats gehad. Tevens zijn studiedagen over de verande-

ringen gehouden. Geen van de scholen heeft nog in deze fase nieuwe leerboeken (methoden) aangeschaft.

In school A en C is men gestart met een inhoudelijke discussie over basisvorming. In school B heeft men eerst organisatorische maatregelen getroffen. Voor alle scholen geldt dat er een proces van schoolwerkplanontwikkeling is begonnen (school B) of is voortgezet (school A en C).

De voorlichting is naar diverse doelgroepen gegaan (intern: docenten en extern: basisscholen). A en C zijn actief in hun voorlichting, terwijl school B dat iets minder is.

School A heeft het gevoel dat de basisvorming goed aansluit bij de onderwijskundige opvattingen in de school. Men ervaart geen sterke tijdsdruk en werkt een over de twee voorbereidingsjaren verdeeld programma van besluitvorming door.

In school B is er sprake van enige gereserveerdheid over de mogelijkheden tot realisering. In school C neemt de schoolleiding de basisvorming serieus. De docenten zijn daar nog op bescheiden wijze betrokken bij de meningsvorming. Sommige aspecten van basisvorming worden ervaren als duidelijke vernieuwingen. In school C is er minder collectief over de basisvorming vergaderd dan in school A en B.

Eind 1992 was dus in geen van de scholen de besluitvorming over de eigen vormgeving van basisvorming afgerond. Als een van de grootste problemen zien school A en C de relatie tussen VBO en basisvorming, terwijl in school B het schoolmanagement als zodanig wordt beschouwd. Voorts ziet men problemen op het gebied van het lesmateriaal (A en C) en de bevoegdheidseisen (C).

4.3 Onderwijsleerproces en leereffecten

In deze fase van het project zijn op klasniveau nog geen onderzoeken uitgevoerd. De instroomgegevens van leerlingen zijn wel verzameld. Vanaf het schooljaar 1993 komen onderwijsleerprocessen bij enkele vakken aan bod. De aandacht gaat dan in het bijzonder uit naar het toepassingsgerichte karakter van basisvorming, de mate van individualisering (heterogeniteit) en schoolgerichtheid, samenwerking en hechtheid. De laatste twee thema's zijn in het voorbereidingsjaar instrumenteel ontwikkeld. De gegevens daarvan volgen hierna.

Het is een bekend gegeven dat leerlingen onderling verschillen in de mate waarin ze zich met de school verbonden voelen. Het is eveneens bekend dat klassen binnen een school zich van elkaar onderscheiden, onder meer in de mate waarin de leerlingen zich op elkaar betrokken voelen. De verbondenheid met de school (schoolgerichtheid) en met de klas (hechtheid in de klas) geeft naar ons idee een indruk van de mate waarin leerlingen zich op school thuisvoelen.

De invoering van basisvorming, in het bijzonder de van beleidswege gewenste fusies om tot meer omvangrijke en meer heterogene scholengemeenschappen te komen, zou tot effect kunnen hebben dat de leerlingen hun klas als gefragmenteerder ervaren (minder hecht) en zich tevens minder verbonden voelen met de school als totaliteit (minder schoolgericht). Moet nu verwacht worden dat een dergelijk effect zonder meer optreedt? Veel zal afhangen van de maatregelen die een school treft om het gevoel van zich verloren wanen in een brede, heterogeen samengestelde, grote scholengemeenschap tegen te gaan. Eén zo'n maatregel zou het zich sterk maken voor samenwerking bij leren (samenwerkingsgerichtheid) kunnen zijn. Immers, samenwerking bij leren beoogt onder meer het gevoel te versterken erbij te horen en scholing waardevol te vinden. Het zijn beide aspecten die bijdragen aan een identificatie met de school (vergelijk het participatie-identificatiemodel van Finn, 1989).

In het Bavo-project gaan we na hoe de integratie van de leerling in de school en in de klas vorm krijgt voorafgaande aan en na de officiële invoering van de basisvorming. In juni 1992 zijn op drie scholen gegevens verzameld over de schoolgerichtheid en hechtheid in het eerste leerjaar van brugklasleerlingen (school A, N=114, school B, N=149, school C, N=146). Over het geheel genomen blijken de drie scholengemeenschappen redelijk overeen te komen in de mate waarin de leerlingen zich thuisvoelen op hun school en klas. Het merendeel van de betrokken leerlingen (70-90%) voelt zich met hun school verbonden (schoolgerichtheid): ze staan positief tegenover de school, hechten aan het goed doorlopen van de school, moeten weinig tot de orde geroepen worden en blijven vrijwel niet zonder toestemming van school weg. De verbondenheid met de klas

(hechtheid in de klas) is op alle drie scholengemeenschappen bij het merendeel van de onderzochte klassen minder dan verwacht kon worden op grond van een norm die gebaseerd is op eerder onderzoek (Cillessen, 1991; Van Boxel, 1993).

Deze bevindingen (de drie scholen verschillen onderling niet sterk in schoolgerichtheid en hechtheid in de klas, de schoolgerichtheid is bij alle drie relatief sterk en de hechtheid in de klas is bij alle drie in veel klassen vrij zwak) staan nog weinig conclusies toe. Het gaat om een eerste meting die betekenis krijgt in de reeks van herhaalde metingen die jaarlijks zullen plaatsvinden tot 1996 op alle drie scholengemeenschappen.

Naast het registreren van het invoeringsproces in het algemeen in de drie scholen richt het Bavo-project zich ook op het klasniveau. Ter toetsing van de vermelde verwachtingen zijn in het vervolg verschillende onderzoeksactiviteiten noodzakelijk. Nader geëxploreerd worden de ontwikkelingen op een drietal vakgebieden (Engels, wiskunde en muzische en beeldende vakken). Voor de komende jaren staan op schoolniveau nog metingen op stapel betreffende

- de opvattingen, problemen (visie) en betrokkenheid van het personeel inzake basisvorming;
- de veranderingscapaciteiten, in het bijzonder de mate van overeenstemming in doelstellingen, de rol van de schoolleider, de communicatie;
- de coördinatiemechanismen, de positie van vaksecties.

Op klasniveau zal het gaan om metingen betreffende

- karakteristieken en effecten van basisvorming in onderwijsleerprocessen en leermiddelen;
- heterogeniteit en individualisering;
- samenwerking, schoolgerichtheid en hechtheid.

5 Perspectief

Internationaal worstelen beleidmakers al vele jaren met vraagstukken rondom de veranderbaarheid van scholen (zie o.m.Fullan, 1991, p. XIII; Van Velzen, Miles, Ekholm, Hameyer & Robin, 1985). Er is sprake van frustraties op tenminste twee niveaus. In de scholen en onder leraren heerst de opvatting dat het beleid maar doorgaat met het ontwikkelen van nieuwe voorstellen, terwijl de actuele veranderingsopdrachten nog niet in praktijk zijn gebracht. Op het niveau van beleidmakers bestaat de gedachte dat onderwijsverbetering te langzaam gaat en dat er nodig een andere koers moet komen.

In ons land ontwikkelt de centrale overheid in de jaren negentig een sterke tendens de beleidsruimte voor scholen te vergroten in de verwachting dat daardoor wenselijke veranderingen beter en sneller tot stand zullen komen. In hoeverre deze verwachting bewaarheid zal worden moet nog blijken. In elk geval betekent meer autonomie voor scholen dat een groot beroep wordt gedaan op eigenschappen die betrekking hebben op capaciteiten van de school als organisatie.

De vraag is of meer autonomie ook leidt tot meer implementatie van veranderingen. Invoering van de basisvorming geeft scholen zeker mogelijkheden om een eigen beleid vorm te geven. Deze vernieuwing kan niet in één schooljaar volledig geïmplementeerd zijn. Het gaat hier in feite niet om eenmalige veranderingen. Basisvorming invoeren betekent een voortdurend proces van ontwikkeling van het onderwijs in een school. Werken aan deze onderwijsveranderingen is werken aan schoolontwikkeling (zie de innovatietheoretische uitgangspunten van dit project).

In deze zienswijze moet niet in de eerste plaats de implementatie van de vernieuwing aandacht krijgen, maar moet de schoolontwikkeling centraal staan. Dat heeft consequenties voor de innovatietheoretische inkadering. Onderzoek en theorieontwikkeling op het gebied van onderwijsinnovatie zijn toe aan een heroriëntatie. We zijn de fase waarin alles draaide om implementatie voorbij. "Beyond implementation leads us to consider more holistic, and organic questions of how individuals and organizations can become better equiped to manage multiple changes as normal fare. Here success is not whether a given innovation is implemented, but whether the basic capacity to deal with change has developped." (Fullan, 1992, p. 113)

Overheidsbeleid dat gericht is op autono-

mievergroting doet in feite een groot beroep op de veranderingscapaciteiten van scholen. Onderzoek laat zien dat op dit moment veel scholen nog onvoldoende toegerust zijn met vermogens om optimaal de beleidsruimte te benutten. Voor innovatieonderzoek betekent dit dat vooral procesonderzoek op langere termijn wenselijk is. Het Bavo-project is vanuit dat perspectief opgezet.

Noten

1 Aan het Bavo-project nemen deel: N. A. J. Lagerweij (projectleider), H. Franssen, R. A. de Jong, G. Kanselaar, J. van der Linden, E. Roelofs, O. Treep, L. Vriens, L. van Wessum.

2 In het onderzoek dat bij al deze projecten verricht is (voor een overzicht zie o.m. Karstanje, 1988; Hofman e.a., 1988a) zijn verschillende min of meer onderling verwante begrippen aan te treffen ter aanduiding van de activiteiten die op schoolniveau uitgevoerd zijn. Men gebruikt – overigens ook voor projecten in het basisonderwijs – termen als 'interne regulering' (Knip & Van der Vegt, 1991), 'schoolontwikkelingsbeleid' (Hofman & Haanstra, 1986), 'probleemoplossend vermogen' (Lagerweij, 1987), 'innovatiecapaciteit of veranderingscapaciteiten van de school' (Van Gennip, 1991), 'interne capaciteiten van scholen' (Hofman & Lugthart, 1991). Dit laatste begrip heeft ook betrekking op de aanwezigheid en de vormgeving van coördinatiemechanismen (in de betekenis van Mintzberg, 1991, p. 118).

3 Enkele veel in internationaal kader geciteerde 'klassieke' gevalstudies over onderwijsinnovatie zijn te vinden bij: Miles, (1964), Gross, Giacquinta, Bernstein (1971) en Smith en Keith (1971). Bekend zijn voorts latere studies van bijvoorbeeld Charters en Pellegrin (1973); Daft en Becker (1978); OECD (1973); Herriot en Gross (1979). In de jaren negentig zijn van belang de studies van onder andere Ramsay en Clark (1990); Louis en Miles (1990); Staessens (1990).

In ons land zijn door verschillende onderzoekers eveneens gevalstudies als onderzoeksmethode gehanteerd (bijvoorbeeld Hofman e.a., 1988b, 1988c; Jehoel, 1984; Van der Krogt, 1983, Van der Krogt & Marx, 1982; Smets, 1985; Imants, 1986; Knip & Van der Vegt, 1991).

Literatuurlijst

Anderson, C. S. (1982). The search for school climate: a review of the research. *Review of Educational Research, 52*, 368-420.

Beem, A. L. (red.) (1988). *Weten, redeneren, raden. Een analystische evaluatie van de basisvorming.* Lisse: Swets & Zeitlinger.

Berg, R. van den, & Vandenberghe, R. (1981). *Onderwijsinnovatie in verschuivend perspectief.* Tilburg: Zwijsen.

Berg, R. van den, & Vandenberghe, R. (1984). *Grootschaligheid in de onderwijsvernieuwing.* Tilburg: Zwijsen.

Boxel, H. van (1993) *Sociometrische status in de adolescentie.* Nijmegen: Universiteitsdrukkerij.

Charters, W., & Pellegrin, R. (1973). Barriers to the innovation process: Four case studies of differentiated staffing. *Educational Administration Quarterly, 9*, 1, 3-14.

Cillessen, T., (1991). *The self-perpetuating nature of children's peer relationships.* Kampen: Mondiss.

Daft, R., & Becker, S. (1978). *The innovative organization: Innovation adoption in school organizations.* New York: Elsevier North-Holland.

Dalin, P. (1992). *How Schools Improve. International Report.* Oslo: IMTEC.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117-142.

Fullan, M. G. (1991). *The New Meaning of Educational Change.* London: Cassell.

Fullan, M. G. (1992). *Successful School Improvement.* Buckingham: Open University Press.

Gennip, H. van (1991). *Veranderingscapaciteiten van basisscholen.* Academisch proefschrift. Nijmegen: ITS.

Giesbers, J. H. G. I., & Krogt, F. J. van der (1990). *Werken met modellen in organisaties: de modellen van Ernst Marx: ontwikkeling, onderzoek en gebruik.* Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.

Gross, N., Giacquinta, J., & Bernstein, M. (1971). *Implementing organizational innovations: a sociological analysis of planned educational change.* New York: Basic Books.

Herriot, R., & Gross, N. (Eds.)(1979). *The dynamics of planned educational change: An analysis of the rural experimental schools program.* Berkeley, CA: McCutchan.

Hofman, R. H., & Haanstra, F. (1986). *Schoolontwikkeling en schoolontwikkelingsbeleid in scholengemeenschappen (eindrapport).* Groningen: RION.

Hofman, R. H., Canne, H. J. de, Olthof, A. L., & Bos, K. T. (1988a). *Onderwijsmanagement en basisvorming. Deelstudie I Literatuurstudie en deskundigenbevraging.* Groningen: RION.

Hofman, R. H., Canne, H. J. de, Olthof, A. L., & Bos, K. T. (1988b). *Onderwijsmanagement en basisvorming. Deelstudie II Innovatieprocessen binnen scholen: een case-study.* Groningen: RION.

Hofman, R. H., Canne, H. J. de, Olthof, A. L., & Bos, K. T. (1988c). *Onderwijsmanagement en basisvorming. Deelstudie III Een smalle en een brede scholengemeenschap op weg naar basisvorming.* Groningen: RION.

Hofman, R. H., & Lugthart, E. (1991). *Interne capaciteiten in het voortgezet onderwijs. (Eindrapport SVO-project 8060).* Groningen: RION, Instituut voor onderwijsonderzoek, Rijksuniversiteit Groningen.

Hopkins, D. (1990). The International School Improvement Project (ISIP), and Effective Schooling: towards a Synthesis. *School Organization, 10,* 195-212.

Imants, J. (1986). *Taakdifferentiatie en besluitvorming in basisscholen.* 's-Gravenhage: SVO.

Jehoel, F. (1984). *Leiding en organisatie van de basisschool.* Den Bosch: Katholiek Pedagogisch Centrum/Malmberg.

Karstanje, P. N. (1988). *Beleidsevaluatie bij controversiële onderwijsvernieuwing.* Amsterdam: Kohnstamm Fonds voor Onderwijsresearch.

Knip, H., & Vegt, R. van der (1991). Differentiated response to a central renewal policy: School management of implementation. *Journal of Educational Policy, 6,* 123-131.

Krogt, F. van der (1983). *Probleemoplossing en beleidsvorming in scholen.* Academisch proefschrift, Rijksuniversiteit Leiden.

Krogt, F. J. van der, & Marx, E. C. H. (1982). *Variaties in schoolorganisaties.* 's-Gravenhage: SVO.

Lagerweij, N. A. J.(1987). Theorie van de onderwijsvernieuwing. In J. A. van Kemenade, N. A. J. Lagerweij, J. M. G. Leune & J. M. M. Ritzen (red.), *Onderwijs bestel en beleid 3: Onderwijs in ontwikkeling* (pp.99-178). Groningen: Wolters-Noordhoff.

Lagerweij, N. A. J. (1993a) Toepassing, Vaardigheden en Samenhang: Basisvorming vereist een schoolconcept. *Handboek basisvorming.* Houten: Van Loghum/ Slaterus.

Lagerweij, N. A. J. (1993b) *Schoolontwikkeling: nieuw perspectief voor onderwijsinnovatie.* Utrecht: Vakgroep onderwijskunde.

Lagerweij, N. A. J., & Voogt, J. C. (1990). Policy making at the school-level: some issues for the 90's. In B. P. M. Creemers & D. Reynolds (Eds.), *School Effectiveness and School Improvement, 1,* 2 (pp. 98-120). Lisse: Swets & Zeitlinger.

Louis, K. Seashore, & Miles, M. B. (1990). *Improving the Urban High School.* New York & London: Teachers College, Columbia University.

Marx, E. (1975). *De organisatie van scholengemeenschappen in onderwijskundige optiek.* Groningen: Tjeenk Willink.

McLaughlin, M. W. (1990). The Rand Change Agent Study Revisited: Macro Perspectives and Micro Realities. *Educational Researcher, 19,* 11-16.

Miles, M. B., & Huberman, A. M. (1984). *Qualitative Data Analysis.* Beverly Hills: SAGE.

Ministerie van Onderwijs en Wetenschappen (1988). *Invoeringsplan basisvorming in het voortgezet onderwijs.* 's-Gravenhage: SDU.

Mintzberg, H. (1991). *Mintzberg over management.* Amsterdam/Antwerpen: Uitgeverij L. J. Veen B.V.

Oakes, J. (1987). *Improving Inner City Schools: Current Directions in Urban District Reform.* Santa Monica, CA: Rand.

Organisation for Economic Cooperation and Development (OECD)(1973). *Case studies of educational innovation: IV. Strategies for innovation in education.* Paris: OECD.

Ramsay, W., Clark, E. E. (1990). *New Ideas for Effective School Improvement.* London: Falmer.

Rosenholtz, S. J. (1985). Effective schools: Interpreting the evidence. *American Journal of Education, 93,* 352-388.

Sleegers, P. & Bergen, T. (1992). De professionele oriëntatie van docenten en de beleidsvoering door scholen. *Velon Tijdschrift voor Lerarenopleiders, 13*(2), 5-9.

Smets, P. (1985). *Coördinatie en management in het voortgezet onderwijs.* Nijmegen: ITS.

Smith, L., & Keith, P. (1971). *Anatomy of educational innovation: An organizational analysis of an elementary school.* New York: Wiley.

Staessens, K. (1990). *De professionele cultuur van basisscholen in vernieuwing (een empirisch onderzoek in VLO-scholen).* Leuven: Katholieke Universiteit, Centrum voor onderwijsbeleid en -vernieuwing.

Stokking, K. M. (1992). *Wijkende condities.* Utrecht: ISOR.

Velzen, W. G. van, Miles, M. B., Ekholm, M., Hameyer, U., & Robin, D. (1985). *Making school improvement work*. Leuven: ACCO.

Voogt, J. C. (1989). *Scholen doorgelicht; een studie over schooldiagnose*. Academisch proefschrift. De Lier: ABC.

Voogt, J. C. (1993). *Onderwijs in beweging. Inforeeks basisvorming 4*. Almere: Procesmanagement basisvorming.

Wetenschappelijke Raad voor het Regeringsbeleid (1986). *Basisvorming in het onderwijs. Rapport aan de regering, no. 27*. 's-Gravenhage.

Yin, R. K. (1989). *Case Study Research. Design and Methods*. Newbury Park/London: SAGE.

Yin, R. K. (1991). The case study method as a tool for doing evaluation. To appear in *Current sociology*. Washington, D.C.: COSMOS Corporation.

Manuscript aanvaard 25-3-1993

Auteur

N. A. J. Lagerweij is als hoogleraar onderwijskunde, in het bijzonder onderwijsinnovatie, verbonden aan de Universiteit Utrecht.

Adres: Universiteit Utrecht, Vakgroep Onderwijskunde, Postbus 80.140, 3508 TC Utrecht

Abstract

Implementation of a core curriculum in Dutch secondary education: the process of initiation.

N. A. J. Lagerweij. Pedagogische Studiën, 1993, 70, 162-179.

This article reports the first findings of an investigation of a current large-scale reform in the Dutch secondary school system. For schools in the secondary schoolsystem the introduction of a core curriculum in 1993 means an extensive combination of far-reaching changes. The new legislation and the policy of the national government give schools some freedom to make their own educational and organizational decisions in various matters according to the core curriculum.

In the Bavo-research-project three schools will be followed throughout five years in their implementation of the core curriculum. These case studies are expected to show the factors which contribute on a school and classroom level to the adoption and implementation as well as the goals of the innovations.

The research is directed to three clusters of variables: schoolcharacteristics, the progress of the process and the content of instructional processes and outcomes. Hypotheses on educational change are formulated for each of these clusters, which will be tested in the case studies. In this article an overview is presented of the data of the adoption period from 1991 until 1993 concerning the beginning phase of the three schools.

Herhaling van woorden en lettergroepen: oefeneffecten bij zwakke lezers

G. Theloosen en W. H. J. van Bon*

Samenvatting

Er werden twee experimenten uitgevoerd die tot hoofddoel hadden na te gaan of jeugdige zwakke lezers een woord gemakkelijker lezen als zij een meerletterig deel van dat woord eerder in andere woorden zijn tegengekomen.

In Experiment 1 kwamen zulke woorddelen of lettergroepen vier of acht maal in andere woorden voor. Deze woorddeel-herhaling bleek geen effect te hebben op het lezen van woorden die die woorddelen bevatten. Herhaling van een heel woord bleek wel een positief effect te hebben op de herkenning van het herhaalde woord, maar een negatief effect op de herkenning van woorden die er voor een deel mee overeenkomen.

In Experiment 2 werd het aantal woorddeel-herhalingen opgevoerd tot zestien. Maar ook onder die conditie werd geen facilitatie geconstateerd op natoetswoorden die het herhaalde woorddeel bevatten. Evenmin werd evidentie gevonden voor de veronderstelde werking van het 'rijm' (het deel van de lettergreep dat bestaat uit de klinker en eventueel erop volgende medeklinkers) als leeseenheid van hogere orde.

Deze experimenten geven niet de verwachte steun aan de veronderstelling dat de lezer decodeereenheden gebruikt op een niveau tussen grafeem en woord. Didactische consequenties van deze bevindingen worden kort besproken.

1 Inleiding

Verschillende leesmodellen veronderstellen dat er twee hoofdroutes of procedures zijn om de uitspraak van een geschreven woord te achterhalen. De ene berust op herkenning van het geschreven woord door gebruik te maken van geheugenopslag van het specifieke letterpatroon van dat woord in het mentale lexicon. Men neemt aan dat herkenning van hoogfrequente woorden voornamelijk langs deze lexicale of directe weg gebeurt. De andere route bestaat uit de samenstelling van het gesproken woord door omzetting van delen van het geschreven woord in hun spraak-equivalenten. Deze indirecte, fonologische werkwijze maakt in beginsel gebruik van de correspondentie tussen grafemen (letters en vaste lettercombinaties zoals *oo*, *ei* en *ch*) en fonemen. Deze routes vinden we ook in het model van Patterson en Morton (1985). Dat model veronderstelt drie procedures. Twee daarvan zijn lexicaal, de derde niet. De niet-lexicale, indirecte procedure gebruikt bij de omzetting eenheden van verschillende grootte: niet alleen grafemen, maar ook 'bodies'. Met 'body' duiden Patterson en Morton aan wat bij anderen (bijvoorbeeld Treiman, 1988) een 'rime' heet: dat deel van de syllabe dat gevormd wordt door de klinker en de medeklinkers die erop volgen. Er zijn echter ook andere opvattingen over de aard van grotere decodeereenheden. Voorbeelden van zulke andere eenheden zijn de 'vocalic center group' (Spoehr & Smith, 1973) of het 'spellingpatroon' (Gibson & Levin, 1975).

Gaan we ervan uit dat in het aanvankelijk leesonderwijs de nadruk ligt op de toepassing van grafeem-foneem-correspondentieregels, dan zou voortschrijdende geoefendheid twee vormen kunnen aannemen die elkaar niet uitsluiten. De ene houdt in dat door herhaalde geslaagde grafeem-foneem-decodering van woorden associaties gevormd worden tussen

* Dit onderzoek kwam tot stand dankzij de welwillende hulp en inzet van leerlingen en leerkrachten van de Prof. Bladergroenschool te Deventer, De Roel van der Neutsschool te Zutphen, de Martinusschool te Millingen aan de Rijn, de Carolusschool te Groesbeek en de Casimirschool te Arnhem.

de letterpatronen van die woorden en hun uitspraak. Er zijn inderdaad aanwijzingen (bijvoorbeeld Reitsma, 1983) dat al vroeg in de leesontwikkeling zulke woordspecifieke, lexicale kennis ontstaat. De andere vorm van leren houdt in dat niet alleen de toepassing van de grafeem-foneem-decodeerprocedure wordt geperfectioneerd (versneld en geautomatiseerd), maar ook dat de indirecte procedure wordt uitgebreid met meerletterige omzettingseenheden boven het grafeemniveau. Het onderzoek dat hierna gerapporteerd zal worden, had tot doel te verifiëren of bij zwakke lezers aanwijzingen te vinden zijn voor het leren van zulke meerletterige eenheden. Daartoe zijn twee trainingsexperimenten gedaan waarin de aanbiedingsfrequentie van lettergroepen gevarieerd werd. Getoetst werd of woorden die frequent aangeboden lettergroepen bevatten, beter gelezen worden dan woorden met minder frequent aangeboden letterclusters. In Experiment 1 wordt herhaling van lettergroepen gecontrasteerd met herhaling van hele woorden. In Experiment 2 is de herhalingsfrequentie verhoogd en worden twee typen lettergroep met elkaar vergeleken.

De veronderstelling dat de beginnende lezer vaardiger wordt in de herkenning van woordsegmenten die groter zijn dan een enkele letter of een enkel grafeem, is niet alleen van theoretisch belang. Zij is ook van praktisch belang voor de ontwikkeling van reguliere en remediële leesleermethoden. Immers, als verwerking van grotere lettergroepen behoort tot de procedures die de gevorderde lezer ten dienste staan, dan zouden leesmethoden de verwerving van die procedure en de daarbij horende kennis dienen te bevorderen. Zo'n gedachtengang lijkt ten grondslag te liggen aan het gebruik in het aanvankelijk leesonderwijs van wisselrijen, dat zijn woordrijen als *boom*, *boos*, *boot*, *boor*, waarin de opeenvolgende woorden een of meer grafemen gemeenschappelijk hebben. De herhaling van het segment *boo* in deze voorbeeldwisselrij zou de vorming van de herkenningseenheid *boo* bevorderen. De effecten van een leestraining als deze zouden niet beperkt blijven tot het in de training gebruikte woordmateriaal, maar zouden zich uitstrekken tot niet-geoefende woorden, doordat de grafeem-foneem-omzetting sneller en accurater wordt en de herkenning van woorden op basis van letter-klankkoppeling plaats maakt voor herkenning van grotere segmenten dan de letter. Dit gebruik van eerder geleerde kennis in nieuwe leersituaties, ook wel transfer genoemd, is volgens Simons (1991) het kernproces van al het leren. Reitsma (1988, 1989) concludeerde echter dat voor beginnende lezers wisselrijen geen voordeel hebben boven willekeurig samengestelde woordrijen. Als het al zo was dat woorden waarvan een meerletterig deel middels wisselrijen geoefend was, beter herkend werden dan woorden zonder zo'n meermaals geoefend cluster, dan werden de herhaalde grafeemclusters zelf niet beter gelezen, terwijl men dat wel moet verwachten als die clusters door de oefening als een herkenningseenheid gaan werken. Ook bij leeszwakke kinderen werd het verondersteld oefeneffect van wisselrijen niet gevonden (Reitsma & Dongelmans, 1988). Waar patroonherhaling wel gepaard ging met betere herkenning van nieuwe woorden die dat patroon bevatten, verklaarde Reitsma dat uit een faciliterend effect dat van de herhaalde - en daardoor beter beschikbare - oefenwoorden uitgaat op de orthografische 'buren'.

Wisselrijen mogen dan niet het verwachte effect hebben, dat betekent niet dat de assumptie waarop hun gebruik gebaseerd is, onjuist is. Reitsma (1989) merkte al op dat de leerlingen bij het lezen van een wisselrij gebruik maken van de overeenkomst tussen de woorden waardoor ze die met minder aandacht lezen dan de woorden in een willekeurige rij. Vandaar ook dat voor de eerste woorden in wisselrijen meer leestijd nodig is dan voor de latere woorden (Reitsma, 1988). Juist deze eigen dynamiek maakt wisselrijen misschien tot een minder effectief middel om meerletterige herkenningseenheden te verwerven. Dan zou een woordpresentatie waarbij niet dezelfde lettergroep voortdurend herhaald wordt, maar waarbij de lettergroepen door elkaar aan de beurt komen, een geschikter middel kunnen zijn om de rol van meerletterige fragmenten in de woordherkenning aan te tonen.

In ons eerste experiment wordt nagegaan of het lezen van verschillende woorden waarin dezelfde meerlettergroepen terugkeren, leidt tot versnelde herkenning van woorden waarvan die lettergroepen deel uitmaken. Dat zou wijzen op het functioneren van die lettergroepen als herkenningseenheden. Niet alleen wor-

den woordfragmenten door middel van verschillende woorden herhaaldelijk aangeboden, ook hele woorden worden herhaald om zo de relatieve effectiviteit van letterclusterherhaling te kunnen bepalen. Vergelijking van deze twee 'herhalingstypen' is relevant omdat in leestrainingen waarin gewerkt werd met losse woorden, naast woordspecifieke effecten zelden transfer naar andere woorden werd geconstateerd (Samuels, Miller & Eisenberg, 1979; Reitsma, 1983; Van Daal, Bakker, Reitsma & Van der Leij, 1986; Van Bon, Van Kessel & Kortenhorst, 1987; Ehri & Wilce, 1983).

De proefpersonen in dit experiment en in het volgende zijn zwakke lezers. Deze keuze heeft een praktische reden. Juist ten aanzien van deze lezers is het immers bij uitstek nuttig te weten wat effectieve leesleeromstandigheden zijn. Een bijkomende reden is dat, als we aannemen dat bij hen de leesontwikkeling op dezelfde manier verloopt als bij normale lezers, maar trager, onderzoek bij hen het voordeel heeft dat zij gemakkelijker 'betrapt' kunnen worden op leesprocessen die aan vroege stadia gebonden zijn; bij normale lezertjes zijn de stadia waarin overwegend fonologisch gedecodeerd wordt vermoedelijk snel voorbij.

Bij de oefening in beide experimenten wordt gebruik gemaakt van kortdurende woord-aanbieding omdat onderzoek van Van den Bosch (1991, hfst. 2) aannemelijk maakt dat herkenning van woordsegmenten bevorderd wordt wanneer woorden maar zeer kort zichtbaar zijn. Het is mogelijk dat als gevolg van zo'n beperkte presentatieduur de lezer gedwongen wordt efficiënter waar te nemen. In het eerste experiment bepaalt de proefpersoon de presentatietijd zelf maar is hij geïnstrueerd die tijd zo kort mogelijk te houden. In het tweede experiment wordt de tijd via een automatische procedure beperkt tot een noodzakelijk minimum.

2 Experiment 1

In dit experiment lazen zwakke lezers gedurende acht sessies een aantal woorden. Sommige woorden lazen ze eenmaal, andere vier-, weer andere achtmaal (*identieke herhaling*). Andere woorden bevatten segmenten die vier- of achtmaal in het oefenmateriaal voorkwamen (*segmentherhaling*). Onmiddellijk na de training werd een leestoets afgenomen. Sommige toetswoorden waren identiek aan de woorden die in de training herhaald waren: identieke herhaling, identieke toetsing (I-I). Andere toetswoorden kwamen in een segment overeen met herhaalde woorden: identieke herhaling, segmenttoetsing (I-S). Weer andere toetswoorden bevatten een segment dat over verschillende woorden herhaald was: segmentherhaling, segmenttoetsing (S-S). De centrale variabelen in dit experiment waren herhalingsfrequentie en herhalingstype (segmentherhaling versus identieke herhaling). Het type segment werd niet systematisch gevarieerd en was geen factor in dit experiment.

Als lezers gebruik weten te maken van segment-identificatie is te verwachten dat de herkenningstijd bij het lezen van nieuwe woorden beïnvloed zal worden door het aantal malen dat een segment van het woord eerder aangeboden is geweest. Te verwachten is dat identieke herhaling tot snellere herkenning van dat woord leidt dan segmentherhaling. Identieke herhaling hoeft echter geen positief effect te hebben op de herkenningstijd van een gelijkend woord. Wanneer men het woord *beer* meerdere malen heeft gelezen, is men dan in het nadeel bij het lezen van *heer*? Aanwijzingen dat iets dergelijks kan plaatsvinden vindt men bij Feustel, Shiffrin en Salasoo (1983). Zij constateerden bij identieke herhaling van zowel woorden als non-woorden dat er meer fouten worden gemaakt tegen orthografisch verwante woorden (orthografische buren). Ook Slowiaczek en Pisoni (1986) vonden nadelen van woordverwantschap in een 'primed' auditieve lexicale-decisietaak. In een lexicale-decisietaak moeten proefpersonen van woorden vaststellen of zij bestaan of niet. Er moesten gesproken woorden beoordeeld worden, en aan elk woord (de 'target') ging een ander gesproken woord vooraf (de 'prime'). Naarmate er meer klankovereenkomst was tussen prime en target (een, twee of drie beginfonemen gelijk) was de beslissing meer vertraagd. "This inhibition, although not evident under all conditions, may be due to competitions among phonologically similar candidates" (blz. 236). Daarentegen vonden Martin en Jensen (1988) in hun lexicale-decisie-experimenten met geschreven woorden negatieve noch positieve ef-

fecten van fonologisch gerelateerde primes.

2.1 Methode

2.1.1 Proefpersonen

De proefpersonen waren 38 leerlingen van twee LOM-scholen, 31 jongens en 7 meisjes, die door hun leerkrachten als zwakke lezers werden beschouwd. Hun gemiddelde score op de Een-Minuut-Test (Brus & Voeten, 1973, vorm A), was 26 (*SD*=6.2). Deze score is gemiddeld voor het begin van groep vier van het basisonderwijs. De gemiddelde leeftijd van de leerlingen was echter aanzienlijk hoger dan in groep vier aan het begin van het schooljaar gebruikelijk is, namelijk 10 jaar (*SD*=15 maanden). Controle op intelligentie vond niet plaats. Aangenomen werd dat de leerlingen, gezien de plaatsing op een LOM-school, beschikten over een 'normale' intelligentie.

2.1.2 Materiaal

In elk van de acht trainingssessies werden 35 eenlettergrepige bestaande woorden gelezen met de consonant/vocaal-structuur CVC, CCVC, CVCC, CCCVC, CVCCC of CCVCC. Alle woorden waren klankzuiver: hun uitspraak was volgens eenvoudige letter-klankomzettingsregels uit hun spelling af te leiden. Hun betekenis was vermoedelijk bij de proefpersonen bekend (zie Tabel 1 voor enkele voorbeelden).

Elk woord dat in de training gebruikt werd, behoorde tot een van vijf herhalingscondities: 1. Het hele woord werd in elk van de acht sessies ter lezing aangeboden (I8). 2. Het hele woord werd in vier van de acht sessies aangeboden (I4). 3. Het woord werd slechts in één sessie gelezen (I1). 4. Een bepaald deel van een woord werd in alle acht sessies (in verschillende woorden) herhaald (S8). 5. Een bepaald segment werd in vier sessies (in verschillende woorden) herhaald (S4). De segmenten bestonden altijd uit meerdere letters. Het waren willekeurig gekozen beginmedeklinkergroepen ('onsets', zie beneden), rimes of beginconsonanten met de vocaal. Het woordmateriaal van een sessie bestond uit tien I8-, vijf I4-, vijf I1-, tien S8- en vijf S4-woorden. Het woordmateriaal voor de gehele training van een groep bestond uit 10 I8-woorden, 10 I4-woorden, 40 I1-woorden, 80 S8-woorden en 40 S4-woorden.

De natoets bestond uit 70 woorden: de 10 woorden die acht keer gelezen waren (I8-I), 10 woorden die in een segment overeenkomen met de I8-woorden (I8-S), 10 woorden die vier keer gelezen waren (I4-I), 10 woorden die in een segment overeenkomen met de I4-woorden (I4-S), 10 woorden die één keer gelezen waren (I1-I), 10 woorden waarvan een segment achtmaal gelezen was (S8-S) en 10 woorden waarvan een segment viermaal herhaald was (S4-S).

Om te voorkomen dat eventuele oefeneffecten gecontamineerd waren met de samenstelling van het toetsmateriaal, werden de proefpersonen over twee groepen verdeeld die in EMT-score, leeftijd en sekse-verhouding nauwelijks van elkaar verschilden. Voor beide groepen proefpersonen werd het woordmate-

Tabel 1

Trainingswoorden en toetswoorden in de condities van Experiment 1, voor groep 1 en 2

		trainingssessie								
conditie	groep	1	2	3	4	5	6	7	8	toetswoord
I8-I	1	peer	peer	peer	peer	peer	peer	peer	peer	peer
I8-S	1									beer
I4-I	1	smal	hek	smal	hek	hek	smal	hek	smal	smal, hek
I4-S	1									dal, heg
S8-S	1	zaal	kaal	staal	baal	maal	kwaal	paal	taal	zaal
S4-S	1	maag	fluit	vraag	fles	flits	vaag	flop	zaag	maag, fluit
I1-I	1	vork	zoet	wijn	rots	verf	hoef	poort	wolf	poort, wolf
I8-I	2	kaal	kaal	kaal	kaal	kaal	kaal	kaal	kaal	kaal
I8-S	2									zaal
I4-I	2	fles	vraag	fles	vraag	vraag	fles	vraag	fles	fles, vraag
I4-S	2									maag, fluit
S8-S	2	beer	peer	speer	leer	meer	heer	keer	teer	beer
S4-S	2	heg	dal	hek	smal	wal	hert	bal	helm	dal, heg
I1-I	2	vork	zoet	wijn	rots	verf	hoef	poort	wolf	poort, wolf

riaal zo samengesteld dat de toetswoorden die bij de ene groep dienden om het effect van identieke herhaling (*peer, peer, ...*) vast te stellen op een woord dat in een segment (*beer*) overeenkomt met de herhaalde woorden, bij de andere groep fungeerden als toetswoorden voor het effect van segmentherhaling (*beer, peer, ...*). Een tekort aan bekende woorden maakte het gewenst bij segmentherhaling de toetswoorden als oefenwoorden op te nemen in de eerste trainingssessies. De I1-woorden waren voor beide subgroepen dezelfde.

2.1.3 Training

De training gebeurde individueel. Op het scherm van een ATARI-ST computer verscheen een fixatiepunt, gevolgd door een woord. Met een druk op een 'muisknop' gaf de leerling te kennen het woord gelezen te hebben (volgens de instructie moest dat 'snel en goed' gebeuren). Onmiddellijk hierop verdween het woord en werd een maskering zichtbaar. De tijd tussen de presentatie van het woord en de druk op de knop werd vastgelegd in honderdsten van een seconde. Nadat de proefpersoon had gezegd wat het woord was, kreeg hij feedback over de snelheid en correctheid van het antwoord. Was een woord fout gelezen dan vertelde de proefleider wat het correcte antwoord was. Het volgende woord werd ingeleid door het fixatiepunt. Voorafgaand aan sessie 1 en sessie 2 werden er tien resp. vijf oefenitems aangeboden.

2.1.4 Toetsing

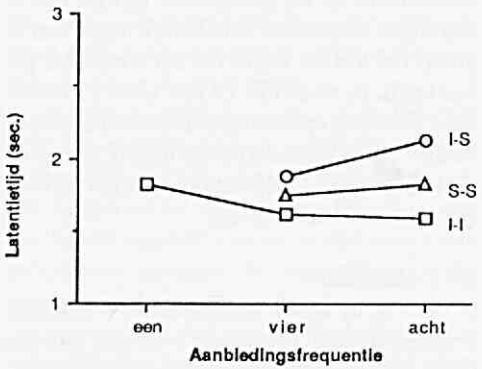
In de natoets, die individueel werd afgenomen, las de leerling het woord dat op het computerscherm gepresenteerd werd en de proefleider drukte op een knop zodra de leerling het woord zei. Ook hier werd in de instructie de nadruk gelegd op 'snel en goed' reageren.

2.1.5 Procedure

De training vond plaats in een periode van acht opeenvolgende schooldagen; één trainingssessie per dag. De toetssessie vond plaats op de eerstvolgende schooldag na de training. De gemiddelde tijd die nodig was om de woorden van een trainingssessie door te nemen bleek ongeveer zeven minuten te zijn. Afname van de toets vroeg ongeveer tien minuten.

2.2 Resultaten

Per proefpersoon werd over de tien toetswoorden waaruit een toetsconditie (I8-I, I8-S, I4-I, enzovoorts) bestond de mediane latentietijd van de goede responsies berekend. Door het gebruik van de mediaan wordt de invloed van uitbijters geminimaliseerd. De data van de twee groepen proefpersonen werden daarbij samengenomen. Het gemiddelde van de latentietijdmedianen per conditie is weergegeven in Figuur 1.



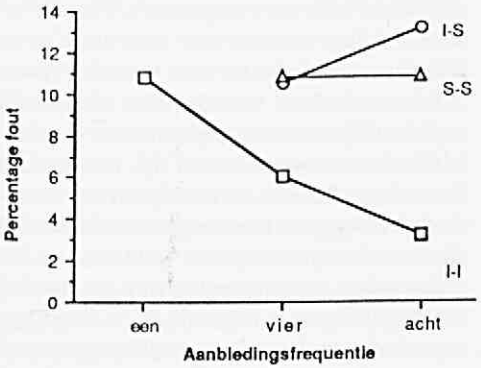
Figuur 1. De gemiddelde latentietijden per conditie (Experiment 1)

Een 3 (materiaal: I-I, I-S, S-S) * 2 (frequentie: 4, 8) ANOVA met herhaalde metingen is uitgevoerd op de mediaantijden; toetsconditie I1-I bleef in deze analyse buiten beschouwing. Er werd zowel een effect van materiaal ($F(2,37)=29.41$, $p<0.01$) als van frequentie ($F(1,37)=9.79$, $p<0.01$) gevonden. I-I-woorden werden sneller herkend dan I-S-woorden ($F(1,37)=46.83$, $p<0.01$). Ook S-S-woorden werden sneller herkend dan I-S-woorden ($F(1,37)=19.44$, $p<0.01$).

De gegevens laten een interactie zien tussen de materiaal- en frequentiecondities ($F(2,37)=5.37$, $p<0.01$). Voor de I-S-woorden is het verschil tussen vier en acht presentaties significant ($F(1,37)=11.06$, $p<0.01$), met langere latentietijden na acht presentaties. Bij de I-I- en de S-S-woorden zien we geen effect van herhalingsfrequentie ($F(1,37)=0.49$ resp. 1.07). Een aanvullende analyse op de I-I-data laat zien dat de afname in latentietijd tussen de frequenties een en vier significant is ($F(1,37)=18.38$, $p<0.01$).

Het gemiddelde foutenpercentage is weer-

gegeven in Figuur 2. Een 3 (materiaal) * 2 (frequentie) ANOVA, met foutenpercentage als afhankelijke variabele, laat zien dat er alleen een hoofdeffect is van materiaalconditie ($F(2,37)=14.55, p<0.01$). Er is noch een hoofdeffect van frequentie ($F<1$), noch een interactie-effect ($F(2,37)=1.92, p>0.10$). Vergelijking van de materiaalcondities I-I en I-S levert een significant verschil op ($F(1,37)=29.05$, $p<0.01$); in de I-I-conditie werden minder vaak fouten gelezen dan in de I-S conditie. Het verschil tussen I-S en S-S is niet significant ($F<1$). Een nieuwe ANOVA laat zien dat de invloed van frequentie (1, 4 en 8) binnen de I-I-conditie significant is ($F(2,37)=9.06$, $p<0.01$). Dit betreft zowel I8-I versus I4-I ($F(1,37)=4.97$, $p<0.05$) als I4-I versus I1-I ($F(1,37)=7.45, p<0.01$). Met het toenemen van de herhalingsfrequentie werden er in de I-I-conditie minder fouten gemaakt.



Figuur 2. Het percentage fouten per conditie (Experiment 1)

2.3 Discussie

Woorden worden het snelst en met de minste fouten herkend na identieke herhaling. Herkenning duurt lang wanneer het woord lijkt op een vaak herhaald woord. Het hier gevonden frequentie-effect bij identieke herhaling (I-I) komt overeen met het in de literatuur gerapporteerde woordfrequentie-effect (bijvoorbeeld Mitchell, 1982). In dit experiment zien we vooral een verschil tussen enerzijds de I1-I-conditie en anderzijds de I4-I- en I8-I-condities. Dat de I4-I-leessnelheid niet verschilt van die van I8-I, en het foutenpercentage wel, suggereert dat veel woordherhaling in deze fase voorlopig nog niet tot verdergaande versnelde herkenning leidt, maar hoogstens tot accurater lezen. Er is (mogelijk tijdelijk) sprake van een bodemeffect in de snelheid. Dat de leessnelheid niet-lineair afneemt met het toenemen van het aantal aanbiedingen vindt in ander onderzoek de nodige steun (Van Bon, e.a. 1987; Van Daal, e.a. 1986). Bij segmentherhaling is in dit experiment geen frequentie-effect gevonden. Voor de afwezigheid van zo'n frequentie-effect is een aantal verklaringen mogelijk. Een verklaring is dat het herhaalde segment niet als eenheid waargenomen en verwerkt is. De situatie is dan te vergelijken met die van vier of acht geheel van elkaar verschillende woorden. Een andere verklaring is dat verwerking van het gepresenteerde woord (ook) leidt tot activering van andere woorden die hetzelfde segment bevatten, bijvoorbeeld verwante woorden die bij eerdere gelegenheden gepresenteerd werden. Deze verklaring lijkt echter minder juist. Men mag dan immers verwachten dat in de S8-S-woorden meer verwarring zal optreden omdat deze immers recenter na elkaar zijn aangeboden. Zowel de latentietijden als de foutenpercentages laten echter geen significante verschillen zien tussen (S8-S) en (S4-S). Ook zijn deze tijden en foutenpercentages gelijk aan die van I1-I zodat met deze woorden niets bijzonders aan de hand lijkt te zijn. Nog een andere verklaring is dat er te weinig herhaald is. Herkenning van segmenten vindt misschien pas plaats wanneer deze vaker dan in dit experiment het geval was, zijn gepresenteerd. Dit is een van de hypothesen die in experiment 2 zal worden getoetst. Hoe het ook zij, vooralsnog moeten we ons beperken tot de constatering dat de in dit onderzoek betrokken proefpersonen geen positieve effecten ondervonden van segmentherhaling.

De vertraging van identieke herhaling op 'segmentverwante' woorden (I8-S) stemt overeen met de resultaten van onderzoek dat al eerder vermeld werd (Feustel, e.a. 1983; Slowiaczek & Pisoni, 1986). Als het woord lijkt op een ander woord dan is het moeilijk te identificeren. In dit verband is ook de constatering van Kusters, Noordman en Van Bon (1986) van belang dat naarmate er meer overeenkomst is tussen een te lezen woord en de foutieve verklanking er minder gecorrigeerd wordt. Zij suggereren dat de deelprocessen die bij indirecte herkenning betrokken zijn nogal eens dezelfde

uitkomst opleveren bij op elkaar gelijkende woorden. De kans op detectie van verschil op grond van de uitkomsten van slechts enkele deelprocessen is dan gering. Voor de proefpersonen in dit experiment zou dit kunnen betekenen dat de uitkomsten van deelprocessen in verband gebracht kunnen worden met meerdere in aanmerking komende woorden. Omdat een vaak herhaald woord weinig informatie nodig heeft om te 'vuren' zullen bij presentatie van een gelijkend ('segmentverwant') woord de uitkomsten van de deelprocessen dit laagdrempelige woord activeren. Nu zijn er twee mogelijkheden. Of de foutieve uitkomst wordt niet als fout herkend, of een controleproces signaleert de fout en er moet opnieuw gelezen worden. Dus zullen er meer fouten en langere herkenningstijden zijn, precies dat wat we bij I8-S zien[1]. Dit zou zich vooral moeten voordoen bij zwakke lezers, omdat zij bij woordherkenning overwegend gebruik maken van een globale waarneming, en minder oog hebben voor details binnen een woord (Reitsma, 1983). Raaymakers (1984) zou in een geval als dit vermoedelijk spreken van 'output-interferentie', waarmee hij bedoelt dat als gevolg van versterkte associaties tussen de gebruikte retrieval-cues en de opgehaalde informatie, het ophalen van andere informatie met behulp van dezelfde retrieval-cues wordt bemoeilijkt.

3 Experiment 2

In Experiment 1 werd geen effect van segmentherhaling gevonden. Experiment 2 heeft tot doel te onderzoeken of segmentherkenning wel plaatsvindt bij een hogere herhalingsfrequentie, en zo ja, of 'rime'-herhaling meer effect heeft dan herhaling van de klinker en wat daaraan voorafgaat ('beginherhaling'). In de inleiding is reeds vermeld dat met name de 'rime' mogelijk een functionele eenheid is waarvan onmiddellijke herkenning faciliterend werkt bij woordherkenning.

Er zijn redenen van taalkundige aard (Trommelen, 1983) om aan de lettergreep een hiërarchische structuur toe te kennen waarin boven het niveau van de afzonderlijke fonemen zich een niveau bevindt waarop de lettergreep is opgedeeld in 'onset' en een 'rime' (in navolging van Trommelen (1989) gebruiken we het Nederlandse 'aanzet' en 'rijm'). Het rijm ('body' in het model van Patterson en Morton) bestaat uit de klinker en eventueel alles dat in de lettergreep er op volgt (zoals *aks* in *straks*); de aanzet is het deel van de lettergreep dat aan het rijm voorafgaat (*str* in *straks*). Aanwijzingen voor de taalpsychologische relevantie van de verdeling in aanzet en rijm zijn bijvoorbeeld gevonden in de systematiek van versprekingen (Treiman, 1988). Ze zijn ook gevonden in het lezen en in aan het lezen verwante vaardigheden. Kirtley, Bryant, McLean en Bradley (1989) lieten zien dat, voorafgaand aan het leren lezen, aanzet-rijm-segmentatie een belangrijke vaardigheid is in de ontwikkeling van het fonologisch bewustzijn. Van Bon en Schreuder (1986) constateerden dat bij fonemische segmentatie fonemen binnen een rijm moeilijker te isoleren zijn dan fonemen waartussen de aanzet-rijm-grens ligt. Onderzoeksresultaten van Treiman en Chafetz (1987) suggereerden dat het voor geoefende lezers eenvoudiger is met aangeboden woordsegmenten woorden samen te stellen wanneer deze segmenten overeenkomen met de aanzet-rijm-verdeling (*tw* en *ist*), dan wanneer de woorden na de klinker zijn opgedeeld (post-vowel-segmentatie: *twi* en *st*). Als het functionele eenheden zijn, redeneerden Treiman en Chafetz voorts, dan zou een verstorend element na de klinker belastender werken dan hetzelfde element vóór de klinker. In het eerste geval zou immers het rijm doorbroken worden, in het tweede geval zou de verstoring samenvallen met de grens tussen aanzet en rijm. In een lexicale-decisietaak vonden zij inderdaad meer vertraging van een vreemd element (twee schuine strepen: //) na de klinker dan van een ervoor. Treiman en Zukowski (1988) vroegen zich af of geoefende lezers bij het uitspreken en spellen van onzinwoorden beïnvloed worden door de uitspraak van andere, reeds bekende woorden. Zij concludeerden dat men bij het assembleren van spraak en spelling grotere eenheden dan de letter gebruikt en dat deze segmenten in de regel corresponderen met de rijmen. Dat vooral rijmen geschikt zouden zijn als te gebruiken segmenten kan te maken hebben met twee omstandigheden. Ten eerste correspondeert het rijm met een natuurlijke eenheid binnen de gesproken taal. De tweede reden is gelegen in de aard van de Engelse orthografie; de medeklinkers achter de

klinkerletter hebben meer invloed op de uitspraak van die klinker dan de voorafgaande medeklinkers (vergelijk *mind* en *find* met *mist* en *fist*). Het ligt daarom voor de hand te veronderstellen, aldus Treiman en Zukowski, dat het onderzoeksresultaat een weerspiegeling is van de kennis die de proefpersonen hebben van de fonologische structuur van de Engelse orthografie. Onderzoek van Bowey (1990), die een priming-procedure gebruikte, ondersteunt de veronderstelling dat orthografische aanzetten en rijmen als visuele woordherkenningseenheden kunnen fungeren. Onderzoek van Kay en Bishop (1987) wijst eveneens op een belangrijke invloed van het rijm in de hardop leestaak, een invloed die een functie is van de aard van het rijm. Is de uitspraak van het rijm van een woord afwijkend van de uitspraak van hetzelfde rijm in veel andere woorden, dan wordt er langzamer gelezen.

Ook bij beginnende lezers werden effecten van aanzetten en rijmen gevonden. Goswami (1986, 1988) kwam tot de conclusie dat analogiegebruik een belangrijke rol speelt in de vroege leesontwikkeling. De kinderen in haar onderzoek maakten meer gebruik van rijmanalogie dan van aanzet-analogie. Wise, Olson en Treiman (1990) lieten beginnende lezers een aantal woorden lezen die in een voorafgaande trainingsfase ofwel op basis van aanzet-rijm ofwel op basis van post-vowel-segmentatie auditief en visueel waren aangeboden. In de testfase moeten dezelfde woorden nogmaals, maar dan zonder de auditieve en visuele hulp, worden gelezen. Het belangrijkste resultaat is dat er meer 'aanzet-rijm-woorden' correct worden gelezen dan 'postvowel woorden'. In onderzoek van Ehri en Robbins (1992) lazen twee in decodeervaardigheid vergelijkbare groepen beginnende lezers, een aantal nieuwe woorden. De analogie-groep had tevoren een aantal analogiewoorden leren lezen die in rijm overeenkwamen met de toetswoorden. De controlegroep had woorden geleerd die in de klinker overeenkwamen met de toetswoorden. In overeenstemming met de verwachting is de bevinding dat de analogie-groep meer toetswoorden correct las dan de controlegroep.

Er is echter ook onderzoek waarin geen evidentie voor aanzet en rijm als functionele eenheden bij het lezen werd gevonden. Van den Bosch (1991) vond met gelijksoortige experimenten als die van Treiman en Chafetz (1987) zowel bij beginnende lezers als bij volwassen lezers geen grotere vertraging ten gevolge van een verbreking van de woordstructuur na de klinker dan van een verbreking vóór de klinker. Wise c.s. (zie Van den Bosch, 1991, blz. 117) slaagden er niet in de uitkomsten van hun eerdere onderzoek te repliceren. Ook blijken wisselrijen waarin het rijm gelijk blijft niet tot betere effecten te leiden dan wisselrijen waarin het beginsegment tot en met de klinker gelijk blijft (Reitsma, 1988; Reitsma & Dongelmans, 1988). In het licht van deze strijdige evidentie is het nuttig de effectiviteit van op aanzet-en-rijm georiënteerde oefeningen te onderzoeken. Om de kans op fonologisch decoderen - en daarmee op segmentverwerking - te maximaliseren, is in dit experiment gebruik gemaakt van pseudowoorden.

3.1 Methode

3.1.1 Proefpersonen

Negentien leerlingen, zeven meisjes en twaalf jongens, van twee LOM-scholen, deden aan het experiment mee. Hun gemiddelde leeftijd was 10;2 jaar (*SD*=11 maanden). De gemiddelde score op de Een-Minuut-Test was 28 (*SD*=6.7), de gemiddelde score voor groep vier van het basisonderwijs. Dit betekent dat proefpersonen een leesachterstand van twee tot drie jaar hadden.

Voor de toewijzing van het woordmateriaal werden de negentien proefpersonen opgedeeld in een groep van tien en een groep van negen.

3.1.2 Materiaal

De kern van het woordmateriaal was een toets die bestond uit 84 éénlettergrepige pseudowoorden. De pseudowoorden die in de training gebruikt werden, zijn op de volgende wijze van het toetsmateriaal afgeleid. Van 36 toetspseudowoorden is in de training het rijmsegment een aantal malen aangeboden in van elkaar verschillende pseudowoorden (twaalf rijmen zijn viermaal, twaalf rijmen achtmaal en twaalf rijmen zijn zestienmaal aangeboden). Van nog eens 36 pseudowoorden werd in de training het beginsegment (eerste medeklinker(s) plus klinker) herhaald, met hetzelfde frequentiepatroon als bij de rijmherhaling. Elke set van 36 toets-pseudowoorden bevatte twaalf

CVC-, zes CVCC-, zes CCVC- en twaalf CCVCC-woorden. Alle woorden werden geschreven met eenletterige klinker- en medeklinkergrafemen. De 36 toets-pseudowoorden waarvan één groep proefpersonen de beginsegmenten in de training herhaald aangeboden kreeg, waren de woorden voor de rijmherhalingsconditie van de andere proefpersonen (en omgekeerd). Naast deze 72 toetspseudowoorden bevatte de toets nog twaalf pseudowoorden waarvan noch het beginsegment, noch het rijm in de training aangeboden werden (vier CVC's, twee CVCC's, twee CCVC's, vier CCVCC's). Deze laatste twaalf woorden vormden de nul-conditie. De woorden van de verschillende herhalingscondities hadden vóór de training gemiddeld dezelfde moeilijkheidsgraad. Op grond van de gegevens uit een vooronderzoek waren pseudowoorden geselecteerd met ongeveer gelijke benoemsnelheid. Tabel 2 geeft een aantal voorbeelden ter verduidelijking. Met name bij de CVC-woorden was het vrijwel uitgesloten alleen maar gebruik te maken van pseudowoorden. Daarom is geprobeerd het aantal bestaande woorden in elke conditie tot een minimum te beperken.

Het aantal grafemen (en letters en fonemen) waaruit de herhaalde fragmenten bestonden, was voor de twee herhalingstypen gelijk: C*VC*-rijmherhaling en *CVC*-beginherhaling, C*VCC*-rijm- en *CCVC*-beginherhaling, CC*VC*-rijm- en *CVCC*-beginherhaling, CC*VCC*-rijm- en *CCVCC*-beginherhaling.

3.1.3 Training

De training gebeurde individueel. Op het scherm van een ATARI-ST computer verscheen een fixatiepunt, gevolgd door een woord. De leerling kreeg de opdracht dat woord zo snel en zo goed mogelijk hardop te lezen. Na een tevoren bepaalde presentatietijd verdween het woord en werd een maskering zichtbaar. Nadat de proefpersoon had gezegd wat het woord was, kreeg hij feedback over de snelheid en correctheid van zijn responsie. Was een woord fout gelezen dan vertelde de proefleider wat het woord was. Het volgende woord werd ingeleid door het fixatiepunt. Vooraf gaand aan sessie 1 en sessie 2 werden er tien resp. vijf oefenitems aangeboden.

De presentatietijd werd als volgt afgestemd op een accuratessecriterium van 75% goede responsies. Als een leerling van de vier soortgelijke woorden (zelfde woordtype (CVC, CVCC, etc.), zelfde herhalingsfrequentie) er minder dan drie goed had, werd de presentatietijd met tien procent verhoogd. Bij drie goede responsies bleef de presentatietijd ongewijzigd en na vier goede responsies werd hij met tien procent verkort. Voor elke leerling werd vóór de eerste trainingssessie per woordtype proefondervindelijk een aangepaste aanvangspresentatietijd bepaald. Bij die gelegenheid werd de proefpersoon ook bekend gemaakt met de apparatuur, de werkwijze en de aard van het woordmateriaal.

3.1.4 Toetsing

De natoets en de follow-up-toets bestond uit het hardop lezen van de hierboven beschreven 84 toetswoorden. In het midden van het computerscherm werd een fixatiesterretje geprojecteerd dat na korte tijd plaatsmaakte voor een toetswoord. De proefpersoon had de opdracht dat woord zo snel en zo nauwkeurig mogelijk

Tabel 2

Trainingswoorden en toetswoorden in de condities van Experiment 2, voor groep 1 en 2

conditie	groep	trainingssessie 1	2	3	4	5	6	7	8	toets-woord
CVC-rijm, 4x	1	von	–	ron	–	–	pon	–	jon	lon
CCVCC-rijm, 16x	1	blals prals	brals knals	frals stals	glals twals	grals vlals	klals vrals	krals zwals	kwals spals	trals
CVC-begin, 4x	2	lof	–	log	–	–	lop	–	lok	lon
CCVCC-begin, 16x	2	trafs tralp	traft tralt	trakt tramp	traks trams	tralf trans	tralg trant	tralk traps	tralm trasp	trals

hardop te lezen. Zodra de voice-key geluid registreerde werd het woord vervangen door een maskering en werd de tijd bepaald die verstreken was vanaf het begin van de presentatie van het toetswoord[2]. Alle proefpersonen kregen de toetswoorden in dezelfde willekeurige volgorde te lezen.

3.1.5 Procedure

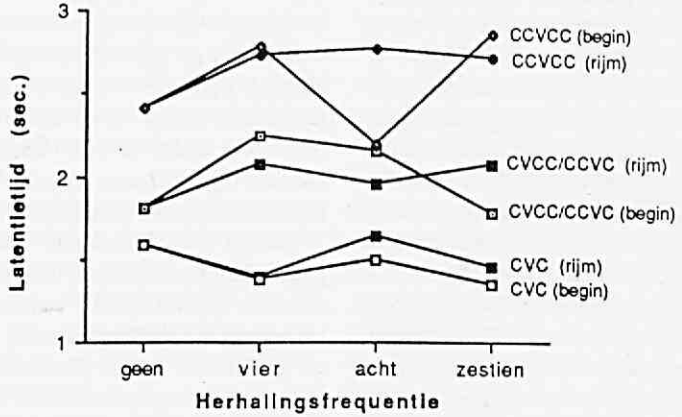
Training en toetsing van een proefpersoon gebeurden in een periode van negen schooldagen. Op elke schooldag was er een trainingssessie. Op de negende schooldag werd de natoets afgenomen, twee weken daarna de follow-up-toets. De trainingssessies duurden ongeveer 12 minuten; de toetsafname duurde gemiddeld 15 minuten.

3.2 Resultaten

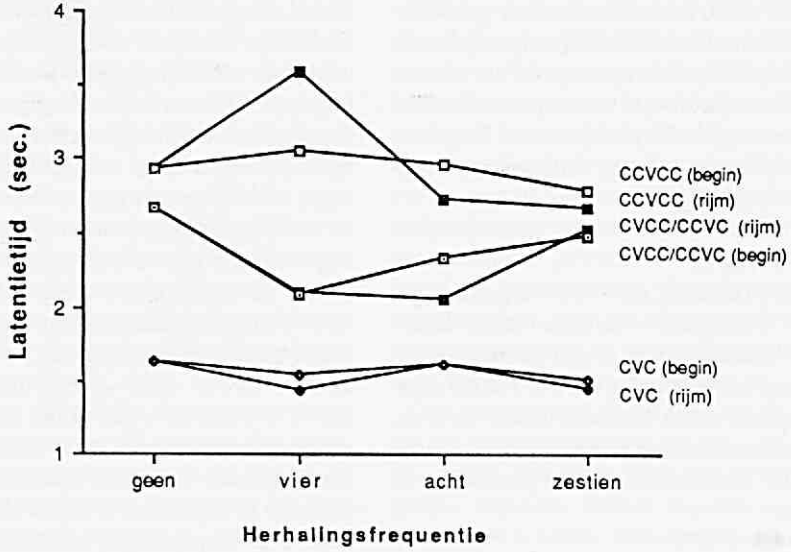
Dertien procent van de responsies op de *natoets* was fout. Over de correct gelezen woorden van elke conditie werd per proefpersoon de mediane latentietijd berekend; de woordtypes CVCC en CCVC werden daarbij samengenomen. In geval de tijdsregistratie foutief was - de 'klok' stopte niet of door een ander geluid dan het hardop gelezen woord - werd de betreffende latentietijd niet meegerekend. Figuur 3 geeft de gemiddelde mediane latentietijden per conditie weer.

Een ANOVA op de mediane latentietijden laat het volgende zien. Van de drie factoren (herhalingstype, herhalingsfrequentie, woordtype) heeft alleen het woordtype (CVC, CVCC/CCVC, CCVCC) een significant effect ($F(2,36)=53.22$, $p<0.01$). Des te korter het woord, des te sneller het gelezen wordt. In deze studie zijn we echter vooral geïnteresseerd in het effect van herhalingsfrequentie en herhalingstype. Daarvan zijn echter geen hoofdeffecten of eerste-orde-interacties gevonden. Het lijkt niet van invloed te zijn of een segment vier-, acht- of zestienmaal herhaald is ($F<1$). Of het herhaalde segment een rijm dan wel een beginsegment is, is eveneens niet van invloed ($F<1$). De significante tweede-orde-interactie van woordtype, herhalingsfrequentie en herhalingstype ($F(4,72)=3.16$, $p<.02$) wordt veroorzaakt door de korte latentietijden van de CCVCC-woorden waarvan het beginsegment achtmaal herhaald was. Aanvullende analyses laten zien dat dat effect niet aan specifieke woorden of segmenten te wijten is. Deze interactie is voorlopig niet te verklaren. Dat de lezers niet hebben geprofiteerd van de herhaalde woordsegmenten blijkt ook uit een vergelijking met de woorden uit de nul-conditie; deze laatste worden niet langzamer gelezen dan de woorden waarvan een segment herhaald werd.

Het percentage leesfouten op de *follow-up-toets* bedroeg twaalf procent. De gemiddelde mediane leestijden van de verschillende condities (over de correcte en bruikbare antwoorden) zijn grafisch weergegeven in Figuur 4. De belangrijkste conclusies uit de gegevens van de natoets gelden ook voor de follow-up-toets. Er is alleen een significant hoofdeffect van het woordtype ($F(2,36)=50.68$, $p<0.01$). Ook hier geldt dat herhaling van woordsegmenten geen effect heeft. Het doet er niet toe of een segment vier-, acht- of zestienmaal herhaald is ($F<1$). Ook hier lijkt het niet van belang te zijn of een



Figuur 3. Gemiddelde mediane latentietijd per conditie (Experiment 2: natoets)



Figuur 4. Gemiddelde mediane latentietijd per conditie (Experiment 2: follow-up-toets)

beginsegment dan wel het rijm herhaald wordt ($F<1$). Een verschil met de natoets is wel het significante interactie-effect van woordtype en herhalingsfrequentie ($F(4,72)=5.71$, $p<0.01$). Figuur 4 laat zien dat de interactie niet eenvoudig te beschrijven is. Het is echter duidelijk dat voor geen van de woordtypen geldt dat meer segmentherhaling leidt tot snellere herkenning.

3.3 Discussie

Het meest betrouwbare effect in Experiment 2 is dat van de woordlengte. Woordlengtegebonden processen spelen klaarblijkelijk nog een belangrijke rol bij de zwakke lezers in dit experiment. Transfer van segmentherhaling in verschillende woorden naar overeenkomstige woorden vond niet plaats. De overtuiging dat leerlingen vaardiger worden in het herkennen van steeds grotere segmenten wordt door de resultaten van dit experiment niet ondersteund. Het maakt daarbij niet uit of een rijm of een beginsegment herhaald wordt. Uitdrukkelijk moet hier echter gesteld worden dat deze resultaten mogelijk gebonden zijn aan de operationaliseringen in dit experiment, d.w.z. bij deze groep zwakke lezers en bij deze opzet.

Dat het rijm bij de identificatie van geschreven woorden een dominante rol zou spelen, zoals in de Angelsaksische literatuur wordt gesuggereerd, vindt hier eveneens geen ondersteuning. De grote overeenstemming tussen de natoets en de follow-up-toets vormt een aanwijzing dat de hier gevonden effecten (voornamelijk afwezigheid van effecten) in hoge mate experimenteel valide zijn.

4 Algemene discussie

De belangrijkste conclusies uit deze twee experimenten laten zich als volgt samenvatten. De lezers in beide experimenten hebben bij het lezen van de toetswoorden geen voordeel gehad van het feit dat ze bij de training woorden gelezen hadden die er op lijken; er vond kennelijk geen transfer plaats. Daarentegen werd de woordherkenningssnelheid wel beïnvloed door het aantal malen dat een woord eerder is gelezen; identieke herhaling leidt tot versnelde herkenning terwijl het lezen van een woord dat orthografisch en fonologisch lijkt op een vaak identiek herhaald woord, bemoeilijkt wordt. Een andere factor die de herkenningssnelheid beïnvloedt is die van woordlengte. Dit suggereert dat bij deze lezers woordlengte-gebonden processen nog een belangrijke rol spelen.

Dat in het Nederlandse taalgebied doorgaans geen aanwijzingen voor de aanzet-rijm-structuur van de syllabe werden gevonden in het leesproces van beginnende, zwakke lezers en gevorderde lezers (Reitsma, 1988, 1989; Reitsma & Dongelmans, 1988; Van den Bosch, 1991) staat in contrast met de onderzoeksresultaten die uit Angelsaksisch onderzoek worden

gerapporteerd. Van den Bosch (1991) voerde ter verklaring aan dat het verschil in correspondentie tussen de orthografie en fonologie in het Engels en het Nederlands hieraan mogelijk debet is. In het Nederlands wordt - anders dan in het Engels - de uitspraak van het klinkergrafeem niet of nauwelijks bepaald door de medeklinkers die er in de lettergreep op volgen.

De theoretische implicatie van onze onderzoeksresultaten is dat het bestaan van functionele eenheden op een niveau tussen letter en woord, die door verschillende woordleesmodellen verondersteld worden, in de empirie van het Nederlandse onderzoek weinig steun vindt. De opvatting dat er een overgangsstadium tussen decoderend lezen en directe herkenning is waarin zulke eenheden een rol spelen, is vooralsnog niet gestaafd.

Overigens moet worden opgemerkt dat de interpretatie van de uitkomsten van het onderzoek naar subsyllabische eenheden niet onweersproken is gebleven. Zo is Seidenberg (1987) van mening dat eventuele segmenteffecten niet noodzakelijk wijzen op het bestaan van functionele eenheden, maar het gevolg kunnen zijn van orthografische redundantie, dus van statistische kenmerken van de verdeling van grafemen binnen woorden die maken dat een grafeem in zekere mate voorspelbaar is uit de aangrenzende grafemen.

Het hier gerapporteerde onderzoek geeft steun aan het onderzoek van Reitsma (1988, 1989; Reitsma & Dongelmans, 1988) naar het effect van wisselrijen. Reitsma concludeert dat het gebruik van wisselrijen niet leidt tot herkenning van lettercombinaties binnen woorden; de 'verkortingshypothese', die stelt dat zoiets wel zou plaatsvinden, wordt niet ondersteund. Zowel bij het gebruik van wisselrijen als bij niet-rijgewijze segmentherhaling zoals in deze experimenten, wordt geen transfer gevonden die wijst op de operatie van grafeemclusters.

In feite was in het hierboven beschreven onderzoek geen sprake van een expliciete training: de proefpersonen werden niet geattendeerd op bepaalde woordstructuren en de herhaalde delen van een woord werden niet geaccentueerd. De taak van de proefpersoon bleef beperkt tot het drukken op een knop of het hardop lezen van woorden; het opsporen van constanten in de woorden behoorde niet tot hun taak. Het is mogelijk dat met hetzelfde woordmateriaal, maar met een andere instructie en een andere presentatie andere effecten gevonden worden; juist zwakke lezers profiteren misschien van een meer expliciete instructie en presentatie. Maar juist indien zo'n instructie die de interne syllabe-structuur benadrukt, de leesprestaties van zwakke lezers zou verbeteren en zou leiden tot grafeemclusterherkenning, zou men bedacht moeten zijn op de mogelijkheid dat dan een bijzondere leesontwikkeling op gang is gebracht waarvan men gewoonlijk - men zie de besproken Nederlandse onderzoeksliteratuur - geen sporen aantreft.

De bevinding uit Experiment 1, dat herhaald lezen van een woord leidt tot betere en snellere herkenning van dat woord, strookt met constateringen die door anderen gedaan werden. Zo vonden Lovett, Warren-Chaplin, Ransby en Borden (1990) een verbetering van de nauwkeurigheid en snelheid waarmee geoefende woorden door kinderen met leesproblemen herkend werden, terwijl het lezen van niet-geoefende woorden niet verbeterde. Dat gold ook voor de oefenconditie waarin minder de nadruk lag op het woordspecifieke letterpatroon en meer op de fonologische werkwijze. In het wisselrijen-onderzoek van Reitsma (1988, 1989; Reitsma & Dongelmans, 1988) werden eveneens hoofdzakelijk leereffecten op geoefende woorden gevonden. Het heeft er dus de schijn van dat de lezer van fonologisch decoderen via grafeemherkenning overgaat op directe woordherkenning, zonder een tussenstadium waarin grafemen als verwerkingseenheden hebben plaatsgemaakt voor lettercombinaties tussen het grafeem- en het woordniveau.

Een en ander suggereert dat in reguliere en remediële leesmethoden de basale grafeemgerichte decodeertraining het best gevolgd kan worden door oefening die gericht is op de herkenning van de woordspecifieke letterpatronen. Er is echter een opmerkelijk verschil tussen de bevindingen uit het wisselrijenonderzoek van Reitsma en uit die van ons Experiment 1. Terwijl wij een negatief effect van woordherhaling op de herkenning van gelijkende woorden vonden, vond Reitsma (zie bijv. 1989, Figuur 1) een positief effect op gelijkende woorden dat hij interpreteert als de faciliterende werking van woordspecifieke

kennis. Dit verschil van negatieve en positieve uitkomsten van woordherhaling hangt wellicht samen met het feit dat in ons Experiment 1 elk herhaald woord het enige met dat specifieke meerletterige patroon was, terwijl in Reitsma's experimenten meerdere woorden met hetzelfde meerletterige patroon herhaald werden. Deze positieve dan wel negatieve bijwerking van woordherhaling in functie van het aantal herhaalde orthografisch verwante woorden zou mogelijk verklaard kunnen worden in termen van spreidende activatie en inhibitie (bijv. McClelland & Rumelhart, 1981). Het lijkt echter hoe dan ook nuttig de hier gesuggereerde samenhang empirisch te verifiëren, alleen al omdat het onderzoeksresultaat richtinggevend zou zijn voor de opzet van oefening in woordspecifieke herkenning. Als onze suggestie juist is, zou het immers de voorkeur verdienen meerdere orthografisch verwante woorden in de oefening op te nemen ten einde generalisatie naar niet-geoefende woorden te bewerkstelligen en om inhibitie-effecten te vermijden.

Noten

1 Helaas zijn de onjuiste antwoorden niet genoteerd. De veronderstelling dat de proefpersonen bij niet-geoefende woorden ernaar zullen tenderen om met het segmentverwante herhaalde woord te antwoorden, kan daardoor niet getoetst worden.

2 Met dank aan N. van Kruysbergen van het Nijmeegs Instituut voor Cognitie-onderzoek en Informatietechnologie voor het vervaardigen van de 'timer-routines'.

Literatuur

Bon, W. H. J. van, Kessel, A. E. G. van & Kortenhorst, E. P. M. (1987). Beïnvloeding van woordherkenningssnelheid door middel van flashcards. In J. Hamers & A. van der Leij (Red.), *Dyslexie 87. Verslag van het congres te Utrecht* (blz. 87-94). Lisse: Swets & Zeitlinger.

Bon, W. H. J. van, & Schreuder, R. (1986). Fonemische analyse: het effect van woordkenmerken. In P. Reitsma, A. G. Bus & W. H. J. van Bon (Red.), *Leren lezen en spellen* (blz. 24-31). Lisse: Swets & Zeitlinger.

Bosch, K. van den (1991). *Poor readers' decoding skills.* Nijmegen: Dissertatie KUN.

Bowey, J. (1990). Orthographic onsets and rimes as functional units of reading. *Memory and Cognition, 18,* 419-427.

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

Daal, V. van, Bakker N., Reitsma, P., & Leij, A. van der (1986). Woordfrequentie, repetitie en ernstige leesproblemen. In P. Reitsma, A. G. Bus & W. H. J. van Bon, (Red.) *Leren lezen en spellen* (blz. 40-43). Lisse: Swets & Zeitlinger.

Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27,* 13-26.

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75,* 3-18.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112,* 309-346.

Gibson, E. J., & Levin, H. (1975). *The psychology of reading.* Cambridge: MIT Press.

Goswami, U. (1986). Children's use of analogy to read: a developmental study. *Journal of Experimental Child Psychology, 42,* 73-83.

Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology, 40a,* 239-268.

Kay, J., & Bishop, D. (1987). Anatomical differences between Nose, Palm, and Foot, or, the body in question: further dissection of the processes of sub-lexical spelling-sound translation. In M. Coltheart (Red.), *Attention and Performance 12* (blz. 449-469). Hillsdale: Lawrence Erlbaum.

Kirtley, C., Bryant, P., MacLean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology, 48,* 224-245.

Kusters, E. D. M., Noordman, L. G. M., & Bon, W. H. J. van (1986). Het benutten van linguïstische informatie voor het corrigeren van fouten door goede en slechte lezers. In P. Reitsma, A. G. Bus & W. H. J. van Bon (Red.), *Leren lezen en spellen* (blz. 64-70). Lisse: Swets & Zeitlinger.

Lovett, M. W., Warren-Chaplin, P. M., Ransby, M., & Borden, S. L. (1990). Training the word recognition skills of reading disabled children: treatment and transfer effects. *Journal of Educational Psychology, 82,* 769-780.

Martin, R. C., & Jensen, C. R. (1988). Phonological priming in the lexical decision task: a failure to replicate. *Memory & Cognition, 16*, 505-521.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. *Psychological Review, 88*, 373-407.

Mitchell D. C. (1982). *The process of reading; A cognitive analysis of fluent reading and learning to read*. Chichester: John Wiley & Sons.

Patterson K. E., & Morton, J. (1985). From orthography to phonology: an attempt at an old interpretation. In K. E. Patterson, J. C. Marshall & M. Coltheart (Red.), *Surface dyslexia* (blz. 335-359). London: Lawrence Erlbaum.

Raaijmakers, J. G. W. (1984). *Psychologie van het geheugen*. Deventer: van Loghum-Slaterus.

Reitsma, P. (1983). *Phonemic and graphemic codes in learning to read*. Amsterdam: Dissertatie VU.

Reitsma, P. (1988). Tussen letter en woord: het effect van oefening met wisselrijen. *Pedagogische Studiën, 65*, 344-357.

Reitsma, P. (1989). Het effect van wisselrijen in een vroege fase van het leren lezen. *Pedagogische Studiën, 66*, 421-428.

Reitsma, P., & Dongelmans, J. (1988). Het effect van oefening met wisselrijen voor leeszwakke kinderen. *Tijdschrift voor Orthopedagogiek, 27*, 248-265.

Samuels, S. J., Miller, N. L., & Eisenberg, P. (1979). Practice effects on the unit of word recognition. *Journal of Educational Psychology, 71*, 514-520.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: access units or orthographic redundancy? In M. Coltheart (Red.), *Attention and Performance 12* (blz. 245-263). Hillsdale: Lawrence Erlbaum.

Simons, P. R. J. (1991). *Transfervermogen*. Nijmegen: Inaugurele rede KUN.

Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition, 14*, 230-237.

Spoehr, K. T., & Smith, E. E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology, 5*, 71-89.

Treiman, R. (1988). The internal structure of the syllable. In G. Carlson & M. Tanenhaus (Red.), *Linguistic structure in language processing* (blz. 27-52). Dordrecht: Reidel.

Treiman, R., & Chafetz, J. (1987). Are there onset- and rime-like units in printed words? In M. Coltheart (Red.), *Attention and Performance 12* (blz. 281-298). Hillsdale: Lawrence Erlbaum

Treiman, R., & Zukowski,.A. (1988). Units in reading and spelling. *Journal of Memory and Language, 27*, 466-477.

Trommelen, M. (1983). *The syllable in Dutch*. Dordrecht: Foris.

Trommelen, M. (1989). De structuur van lettergrepen. In M. P. R. van den Broecke (Red.), *Ter sprake: spraak als betekenisvol geluid in 36 thematische hoofdstukken* (blz. 375-382). Dordrecht: Foris.

Wise, B. W., Olson, R. K., & Treiman, R. (1990). Subsyllabic units in computerized reading instruction: onset-rime vs. postvowel segmentation. *Journal of Experimental Child Psychology, 49*, 1-19.

Manuscript aanvaard 1-3-1993

Auteurs

G. Theloosen, Orthopedagoog, Tilburgse Orthopedagogen Praktijk

W. H. J. van Bon, Universitair Hoofddocent, Instituut voor Orthopedagogiek, KUN

Correspondentie-adres: Dr. W. H. J. van Bon, Instituut voor Orthopedagogiek, Katholieke Universiteit Nijmegen, Postbus 9103, 6500 HD Nijmegen.

Abstract

Repetition of words and letter clusters: practice effects in poor readers

G. Theloosen & W. H. J. van Bon. Pedagogische Studiën, 1993, 70, 180-194.

Two experiments were done in order to verify whether a word is easier to read for young poor readers if a letter cluster of that word has been repeatedly presented in other words.

In Experiment 1, letter clusters were seen four or eight times as parts of other words presented in preceding sessions. This letter cluster repetition appeared to have no effect on the identification of words containing such clusters. Repetition of an en-

tire word had a positive effect on the recognition of the repeated word, but a negative effect on the recognition of similar words.

In Experiment 2, the number of cluster repetitions was increased to 16. But even under that condition no facilitation was found on the reading of words containing the repeated clusters. Also no evidence was found for the supposed operation of the rime (the part of the syllable consisting of the vowel and any following consonant(s)) as a higher level reading unit.

The experiments do not support the hypothesis that readers use decoding units at a level between grapheme and word.

The didactic consequences of these findings are discussed.

De invloed van het gebruikte leerboek op het maken van specifieke fouten bij wiskundeproblemen*

J. Chr. Perrenet

Samenvatting

Wiskundemethoden verschillen in didactische opzet. In onderzoek van De Leeuw, Meijer, Perrenet en Groen (1988) met een zogeheten transfertest werd tussen drie methoden geen verschil gevonden op prestaties van vwo-leerlingen aan het einde van het derde leerjaar. Dit artikel beschrijft een nadere analyse van dat onderzoeksmateriaal op het voorkomen van enkele specifieke fouten in de schriftelijke uitwerkingen door de leerlingen van de transfertestopgaven. t methodekenmerken worden hypothesen afleid betreffende twee typen fouten: rekenuten en wiskundetaal-fouten (fouten in notaen terminologie). Bij beide fouttypen wordt controle voor de invloed van de factor wisndige vaardigheid en de factor leraar geen rschil gevonden. De leerlingen van de meode met vanaf het begin de meeste nadruk op mele wiskundetaal maken tegen de verichting in meer wiskundetaal-fouten, maar t verschil is niet significant. De rol van formele skundetaal in het wiskunde-onderwijs wordt sproken in relatie tot het nieuwe wiskunderplan voor de onderbouw.

Inleiding

Onderzoekers verschillen van mening over het effect van de gebruikte methode op de prestaties van de leerlingen. Binnen het domein van rekenen en wiskunde zien we bijvoorbeeld in Pedagogische Studiën de kritiek van Van den Heuvel-Panhuizen (1990) op methodenonderzoek van Harskamp en Suhre. Zij rapporteren (bijvoorbeeld Harskamp, 1988), dat er globaal slechts marginale verschillen zijn in toetsprestaties tussen leerlingen van de basisschool, die tien verschillende reken/wiskundemethoden hebben gebruikt. Wel is er een samenhang tussen de mate van aandacht voor een bepaald leerstofonderdeel en de prestaties op dat onderdeel, de zogenaamde analogie van Walker en Schaffarzick (1974): methoden doen het beter met betrekking tot die leerstofonderdelen in de toetsen die in de betreffende methode de meeste aandacht krijgen. Van den Heuvel-Panhuizen uit kritiek op uitgevoerde leerstofanalyse en de implementatiecriteria van Harskamp en Suhre, maar wijst vooral ook op het belang van de op het eerste gezicht zo vanzelfsprekende analogie van Walker en Schaffarzick. Kennelijk kunnen leerboeken gezien worden als een krachtig middel om de leerprestaties van leerlingen te beïnvloeden. In Gravemeijer, Van den Heuvel-Panhuizen en Van der Ploeg (1990) wordt gesteld dat verschillende leermethoden leiden tot verschillende prestatieprofielen. Voor het vaststellen van die verschillen zijn echter wel verfijnde toetsen nodig, waarop de verschillen tot uiting kunnen komen. Globale totaalscores zijn daartoe als meting te grof.

Een alternatieve aanpak werd gehanteerd door De Leeuw, Meijer, Perrenet en Groen (1988) bij de vergelijking van de prestaties van leerlingen van drie verschillende wiskundemethoden. Naast een toets met standaardopgaven construeerden zij een zogenaamde transfertest. De toetsopgaven stegen qua moeilijkheid uit boven het niveau van de opgaven in de gebruikte leerboeken en vereisten toepassing van wiskundige kennis en vaardigheden in nieuwe situaties. Bij de opgaven werd optionele hulp geboden in de vorm van hints. De hoeveelheid gebruikte hulp werd in de score verdisconteerd. In Pedagogische Studiën werd eerder over dit onderzoek geschreven (Meijer, Perrenet & Riemersma, 1988). De resultaten

* Met dank aan W. Groen voor zijn commentaar op een eerdere versie van dit artikel en aan J. Meijer ook voor zijn methodologische adviezen.

(zie ook De Leeuw e.a., 1988) gaven aan dat er geen verschil is op de transfertestscore, wanneer werd gecorrigeerd voor het niveau van wiskundige vaardigheid (de vaardigheid in het oplossen van standaardproblemen[1]). Ook op dat niveau van wiskundige vaardigheid is er geen methode-effect: het effect van de factor school is veel belangrijker. Een andere onderzoeker, Van Streun (1990), benadrukte juist de invloed van het (wiskunde-)leerboek. Behalve eigen onderzoek (Van Streun, 1989) voerde hij hiervoor ook de resultaten van het zojuist genoemde transfertestproject aan, die hij op een andere wijze interpreteerde. Meijer, Perrenet en Groen (1992) bestreden deze interpretatie en besloten met de opmerking, dat het gebruikte leerboek zeker enige invloed zal hebben, maar dat die invloed in de regel overschaduwd wordt door leerling- en groeperingskenmerken.

Dit artikel beschrijft een nadere analyse op het materiaal van de transfertest, de schriftelijke uitwerkingen van de opgaven door de leerlingen. Daarbij is nagegaan in hoeverre het voorkomen van bepaalde specifieke fouten is toe te schrijven aan het al dan niet gebruiken van een bepaald leerboek. Op grond van de leergangprofielen die door analyse binnen het transfertestproject werden opgeleverd, worden hypothesen opgesteld betreffende de mate van voorkomen van bepaalde fouten bij bepaalde groepen leerlingen. Net als bij het transfertestonderzoek dient wiskundige vaardigheid bij deze hypothesen als controle variabele. In een latere analyse wordt ook de leraar als factor opgenomen. De resultaten worden gerelateerd aan de aanstaande leerplanhervorming bij wiskunde in de onderbouw van het voortgezet onderwijs.

2 Methode

2.1 Leerganganalyse

De drie onderzochte methoden - Sigma, Getal & Ruimte en Moderne Wiskunde[2] - hadden de naam te verschillen in de mate, waarin nadruk gelegd wordt op het oefenen van bepaalde strakke oplossingswijzen of juist het stimuleren van en vergelijken van variatie in aanpak. Hoe dwingerder het cognitieve proces behorend bij een oplossingsmethode is voorgeschreven, des te meer algoritmisch de methode genoemd kan worden (De Leeuw, 1979)[3]. Het algoritmisch karakter van een methode werd in het transfertestproject bepaald aan de hand van a) de mate waarin oplossingsmethoden door herhaling geoefend worden, b) de mate van volledigheid en herkenbaarheid van uitgewerkte voorbeelden en c) de mate waarin bij uitgewerkte voorbeelden de betreffende oplossingsmethode dwingend wordt voorgeschreven. Uit de analyse kwam Moderne Wiskunde duidelijk als de minst algoritmische methode naar voren en Sigma als de meest algoritmische methode. Een tweede kenmerk is de mate van koppeling aan de realiteit, de mate waarin leerstof en opgaven in een realistische context worden aangeboden. Dit aspect bleek in Moderne Wiskunde veel meer vertegenwoordigd dan in Getal & Ruimte en Sigma. Een derde kenmerk is de mate van gebruik van wiskundige taal: formele terminologie en notaties. Getal & Ruimte kwam naar voren als de methode, waarin - meer dan in Sigma en vooral meer dan in Moderne Wiskunde - daaraan aandacht werd besteed. Deze wiskundige taal was in de methoden Sigma en Moderne Wiskunde niet geheel afwezig, maar kwam pas later en minder nadrukkelijk naar voren. Een meer uitgebreide beschrijving van de leerganganalyse en de resultaten is te vinden in De Leeuw e.a. (1988).

In grote lijnen bleek het ook mogelijk de drie methoden te plaatsen binnen de categorisering die Treffers (1987) geeft voor reken- en wiskundemethoden. In Meijer (1990) wordt op die wijze Sigma mechanistisch genoemd, Getal & Ruimte structuralistisch en Moderne Wiskunde realistisch. Mechanistisch wil dan zeggen: nadruk op oefenen van vaardigheden, weinig aandacht voor inzicht, noch voor toepassingen in de realiteit. Structuralistisch wil zeggen: gericht op inzicht in de structuur van het vak, achteraf beperkte toepassingen in de realiteit. En realistisch wil zeggen: de realiteit zowel als bron en toepassingsgebied voor het wiskunde leren, gericht op inzicht in de structuur van het vak gekoppeld aan de realiteit.

2.2 Meetinstrumenten

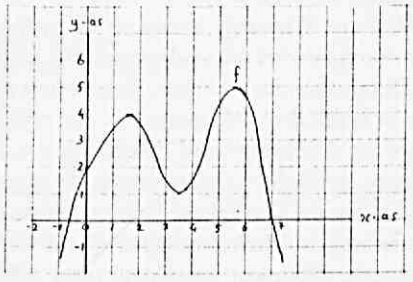
De transfertest bestond uit 15 opgaven, waarvan elke leerling er vijf kreeg aangeboden. De opgaven vereisten gebruik van kennis uit het leerstofdomein functies in nieuwe situaties. Twee voorbeelden zijn gegeven in Figuur 1 en

Figuur 2. Bij de bepaling of een antwoord goed of fout was, werd binnen het transfertestonderzoek vooral gelet op begripsfouten in tegenstelling tot meer technische fouten, die in het algemeen binnen het wiskunde-onderwijs als relatief minder ernstig beschouwd worden. De technische fouten hadden dus geen invloed op de transfertestscore. Op de transfertest werd geen methode-effect gevonden (De Leeuw e.a., 1988). Binnen de categorie van de technische fouten kon een eventueel effect echter niet tot uitdrukking komen. Mede daarom werd deze soort fouten gekozen voor het nadere onderzoek in deze studie naar methode-effecten. De tweede reden is, dat juist op door docenten als minder zwaar beoordeelde fouten de kans groter is eventueel een methode-effect te vinden. Immers: methoden, die zwaardere fouten bij hun leerlingen veroorzaken, zullen zich niet lang handhaven. Het gaat bij de onderzochte methoden om gevestigde methoden, die zich in de praktijk juist als bruikbaar hebben bewezen.

Met technische fouten bedoelen we fouten bij het gebruik van wiskundige taal (wiskundetaal-fouten) en rekenfouten.

Onder *wiskundetaal-fouten* verstaan we alle afwijkingen van het wiskundig taalgebruik, met name notaties en terminologie zoals gehanteerd in de drie leerboeken. We geven een voorbeeld aan de hand van de opgave in Figuur 1.

Hieronder zie je de grafiek van de functie f.



Als de vergelijking f(x)=p vier oplossingen heeft, tussen welke waarden moet het getal p dan liggen?

Figuur 1. Voorbeeldopgave bij wiskundetaal-fouten

Een correcte formele notatie van het eindantwoord is $1<p<4$ of $p\in<1,4>$. Voorbeelden van onjuiste notaties zijn $4<p<1$ en $p=<1,4>$. Een antwoord als $-0.7<p<+7$ duidt op een begripsfout - er is afgelezen op de horizontale in plaats van op de verticale as - maar de notatie is correct. Een voorbeeld van onjuiste terminologie is het verwarren van termen zoals vierhoek en vierkant bij een andere opgave van de transfertest. Er zijn vijf opgaven, waarin wiskundetaal-fouten worden gemaakt. In de analyse zijn die leerlingen betrokken die minstens twee van die vijf opgaven hebben gemaakt. Als maat voor het voorkomen van fouten in notatie en terminologie is genomen het aantal door de leerling gemaakte opgaven met dergelijke fouten ten opzichte van het aantal door de leerling gemaakte opgaven, waarin binnen de gehele groep dergelijke fouten voorkomen, uitgedrukt in een percentage. Een leerling bijvoorbeeld, die drie opgaven heeft gemaakt, waarin binnen de gehele groep wiskundetaal-fouten voorkomen, en bij twee van die opgaven dergelijke fouten maakt, krijgt een score van 67.

Onder *rekenfouten* verstaan we fouten bij de elementaire rekenoperaties als optellen, aftrekken, vermenigvuldigen en delen en ook slordigheden bij het overnemen van gegevens. Als voorbeelden geven we de antwoorden $v=56+12t$ en $v=81+12t$ bij de opgave van Figuur 2. Beide antwoorden zijn ontstaan uit een fout bij het vermenigvuldigen van 9 met 8. Er zijn vier opgaven, waarin rekenfouten voorkomen, waarvan de leerlingen er maximaal twee gemaakt kunnen hebben. Als maat voor het voorkomen van rekenfouten wordt weer genomen het aantal opgaven met een rekenfout ten opzichte van het aantal gemaakte relevante opgaven, net als bij de wiskundetaal-fouten.

Naast de beide effectmaten nemen we twee controlevariabelen op, *wiskundige vaardigheid* en de factor *leraar*. De leraar is opgenomen om een controle uit te oefenen voor de effecten van groeperingsfactoren. De wiskundige vaardigheid dient als alternatief voor een bepaling van het beginniveau (vóór het gebruik van de methoden), dat niet werd gemeten. De test voor wiskundige vaardigheid bestond uit 19 - met 0 of 1 gescoorde - items over het leerstofdomein functies, voorwaardelijke kennis om de transfertest te kunnen maken. De items waren geselecteerd en aangepast

Een meisje werkt 's zaterdags als kassière in een supermarkt. Zij verdient daar f 9,– per uur; voor overwerk krijgt ze f 12,– per uur. De overwerktijd gaat in, nadat er 8 uren zijn gewerkt. Ze moet elke zaterdag overwerken, maar niet altijd even lang.
We noemen v het *totale* aantal guldens dat ze verdient op een bepaalde zaterdag en t het aantal uur *overwerk* van die dag.
Bedenk nu een formule, waarmee je v kunt berekenen als je t weet.

Figuur 2. Voorbeeldopgave bij rekenfouten

op basis van de leerdoelgerichte toetsen van het CITO (1982, 1986). Binnen het tranfertestproject (De Leeuw e.a., 1988) werd voor deze test een redelijke betrouwbaarheid vastgesteld (Cronbachs alpha = .70).

2.3 Steekproef

De transfertest werd afgenomen bij 325 leerlingen uit de derde klas van het vwo, verdeeld over 8 scholen, 10 leraren en 12 klassen. De verdeling is weergegeven in Figuur 3. Bij de methode Moderne Wiskunde bijvoorbeeld waren twee scholen in het onderzoek betrokken; bij de ene school nam één leraar deel met twee klassen, bij de andere school twee leraren met ieder één klas. Per klas waren er 22 tot 30 leerlingen.

Elke leerling kreeg een subset van vijf opgaven. Bij de genoemde typen fouten waartoe alleen bepaalde opgaven aanleiding geven, is dus telkens slechts een deel van de leerlingen relevant. Van de leerlingen, waarbij ook de wiskundige vaardigheidstest werd afgenomen, kon bij 177 het percentage wiskundetaalfouten vastgesteld worden en bij 261 het percentage rekenfouten.

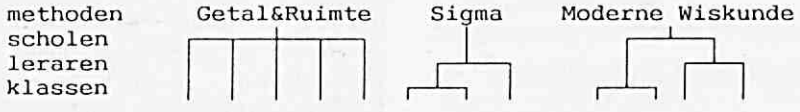
Niet elke leerling kreeg dezelfde opgaven aangeboden, maar per klas (respectievelijk leraar, school, methode) was dat wel het geval (toetsen met chi-kwadraat gaf aan, dat afwijkingen van een overeenkomstige verdeling van opgaven aan toeval konden worden toegeschreven, vrijwel altijd met een kans groter dan 95%). Dat betekent dat uitspraken over gemiddelden op het niveau van klas, leraar, school en methode goed mogelijk zijn. Uitspraken over de mate van fouten van een individuele leerling zijn minder goed mogelijk, aangezien bij beide typen fouten de diverse opgaven bleken te verschillen in de mate waarin ze fouten genereren, zoals bleek uit een chi-kwadraat toets.

2.4 Hypothesen

De hypothesen worden afgeleid uit de leerboekkenmerken zoals vermeld in paragraaf 2.1 en in overeenstemming met de analogie van Walker en Schaffarzick, beschreven in paragraaf 1.

Getal & Ruimte viel op, in vergelijking met Sigma en Moderne Wiskunde, door het vroegtijdige en intensieve formele taalgebruik. Nadrukkelijk zijn wiskundige notaties en termen vanaf de brugklas aanwezig. We poneren de hypothese dat bij de leerlingen die Getal & Ruimte gebruiken gemiddeld minder fouten in terminologie en notatie voorkomen dan bij de leerlingen die één der beide andere methoden gebruiken.

Moderne Wiskunde kwam als de methode naar voren met het minst algoritmische karakter. De opgave-uitwerkingen zijn het minst precies; er is minder vaak sprake van een eenduidig bepaald eindantwoord op een opgave. De mindere aandacht voor precisie in de uitwerkingen zou tot slordigheid kunnen leiden. Te verwachten is dat de leerlingen die Moderne Wiskunde gebruiken gemiddeld meer rekenfouten zullen maken dan de leerlingen van de



Figuur 3. Verdeling van scholen, leraren en klassen per methode

andere methoden.

In beide hypothesen worden dus de resultaten van de leerlingen van een bepaalde methode vergeleken met de resultaten van de leerlingen van beide andere methoden. Voorspeld wordt dat bij Getal&Ruimte de minste wiskundetaal-fouten voorkomen en bij Moderne Wiskunde de meeste rekenfouten. De hypothese met betrekking tot de rekenfouten moet gezien worden als een exploratieve hypothese. Bij het toetsen van de hypothesen en verdere analyse wordt gebruik gemaakt van het programma SPSS/PC+ (SPSS Inc., 1988).

3 Resultaten

Rekenfouten worden in veel mindere mate gemaakt dan wiskundetaal-fouten. De gemiddelde percentages zijn 15.37 voor de wiskundetaal-fouten en 6.32 voor de rekenfouten (zie Tabel 1 en 2). Onder de wiskundetaalfouten treffen we veel onjuiste notaties aan voor de uitdrukking 'p tussen 1 en 4'. Naast de voorbeelden, gegeven in 2.1, noemen we 'p tussen <1,4>', een onjuiste combinatie van formele en informele taal. Verder zien we problemen met de notatie van een functie (fx i.p.v. f(x)). Naast het voorbeeld van de terminologiefout in 2.1 zien we ook ongelukkige aanduidingen van de term richtingscoëfficient. Duidend op de twee rechten, voorgesteld door de vergelijkingen y=x+3 en y=x+1000, wordt bijvoorbeeld gezegd dat 'de x-en hetzelfde zijn'; bedoeld is: 'de richtingscoëfficienten zijn hetzelfde'. Bij de weinige rekenfouten gaat het voornamelijk om vermenigvuldigingsfouten (zie de voorbeelden in 2.1) en verder om vergissingen bij grote getallen (100 of 10.000 in plaats van 1000).

In Tabel 1 geven we een overzicht van de mate waarin de wiskundetaal-fouten worden gemaakt door de leerlingen van Getal&Ruimte in vergelijking met die van de beide andere methoden. Tevens zijn de resultaten voor wiskundige vaardigheid vermeld.

In Tabel 2 staan de gegevens voor de rekenfouten, waarbij in overeenstemming met de hypothese de methoden op een andere wijze gegroepeerd zijn: nu is Moderne Wiskunde tegenover beide andere methoden geplaatst. Ook voor deze groep leerlingen zijn de gegevens voor wiskundige vaardigheid aangegeven.

We zien dat bij de rekenfouten de gemiddelden verschillen in de voorspelde richting. Tabel 2 toont, dat bij de leerlingen van Moderne Wiskunde meer rekenfouten worden gemaakt

Tabel 1

Overzicht percentage wiskundetaal-fouten en niveau wiskundige vaardigheid bij gebruikers van Getal&Ruimte versus Moderne Wiskunde/Sigma

methode	aantal lln.	percentage wiskundetaal-fouten		niveau wiskundige vaardigheid	
		gemid.	*st.dev.*	*gemid.*	*st.dev.*
Getal & Ruimte	75	20.11	28.30	10.24	3.44
Mod. Wisk. of Sigma	102	11.89	21.96	10.74	3.24
Totaal	177	15.37	25.10	10.53	3.32

Tabel 2

Overzicht percentage rekenfouten en niveau wiskundige vaardigheid bij gebruikers van Moderne Wiskunde versus Getal&Ruimte/Sigma

methode	aantal lln.	percentage rekenfouten		niveau wiskundige vaardigheid	
		gemid.	*st.dev.*	*gemid.*	*st.dev.*
Moderne Wiskunde	88	9.09	26.85	11.19	2.96
Getal & Ruimte of Sigma	173	4.91	19.93	9.80	3.46
Totaal	261	6.32	22.54	10.27	3.36

dan bij de leerlingen van de andere methoden (Getal & Ruimte of Sigma). Tabel 1 laat echter zien, dat bij de leerlingen van Getal & Ruimte meer wiskundetaal-fouten worden gemaakt dan bij de leerlingen van Moderne Wiskunde of Sigma en dat is tegengesteld aan de voorspelling. Ten slotte valt in beide tabellen op, dat de standaarddeviaties bij de gemiddelde foutenpercentages relatief groot zijn.

Om de hypothesen te toetsen moeten we corrigeren voor het niveau van wiskundige vaardigheid. Covariantie-analyse met wiskundige vaardigheid als covariaat resulteert bij de rekenfouten in een niet significant verschil (p=.108; F=2.59 bij vrijheidsgraden 1 en 258). We vinden dus geen methode-effect op de mate waarin rekenfouten worden gemaakt. Voor de wiskundetaal-fouten blijkt dat er, ook na controle voor niveau van wiskundige vaardigheid, geenszins sprake is van een kleinere mate van voorkomen van wiskundetaal-fouten bij de leerlingen van Getal & Ruimte in vergelijking met die van Moderne Wiskunde en Sigma. Eerder het tegendeel lijkt het geval. Er is een trend tegen de voorspelde richting in (p=.043; F=4.23 bij vrijheidsgraden 1 en 174)[4].

Vervolgens nemen we de leraar als extra controlefactor op. Tabel 3 presenteert de gegevens voor de wiskundetaal-fouten en wiskundige vaardigheid gegroepeerd per leraar.

Bij de analyse moet er rekening mee worden gehouden dat het gaat om een genest design met de factor leraar binnen de factor methode: eenzelfde methode wordt wel door meerdere leraren gebruikt, maar eenzelfde leraar gebruikt slechts één methode (zie ook Figuur 3). Simultane toetsing van de effecten van leraar en methode met daarbij wiskundige vaardigheid als covariaat resulteert opnieuw in een trend in de richting tegengesteld aan de voorspelling (p=.075; F=4.35 bij vrijheidsgraden 1 en 7). De hypothese, dat de leerlingen van Getal & Ruimte minder wiskundetaal-fouten maken, moet verworpen worden. Interessant is het daarbij nog te vermelden dat de leerlingen van Moderne Wiskunde en Sigma wel in gelijke mate kiezen voor formeel taalgebruik (bijvoorbeeld: $1<p<4$ of $p \in <1,4>$) of informeel taalgebruik (bijvoorbeeld: p tussen 1 en 4) als de leerlingen van Getal & Ruimte.

Een zelfde analyse bij de rekenfouten laat zien dat er ook in dat geval geen sprake is van een methode-effect. Bij controle voor de effecten van wiskundige vaardigheid en de leraar is het verschil in de voorspelde richting niet significant (p=.129; F=2.97 bij vrijheidsgraden 1 en 7).

4 Conclusies en discussie

We vatten de gevonden resultaten samen.

Na controle voor het niveau van wiskundige vaardigheid en de factor leraar is er geen verschil tussen methoden ten aanzien van het voorkomen van rekenfouten. Wel is er een trend in de voorspelde richting. Ook ten aanzien van wiskundetaal-fouten is er geen verschil na controle voor wiskundige vaardigheid en de factor leraar. Nu is er echter een trend te-

Tabel 3

Overzicht percentage wiskundetaal-fouten en niveau wiskundige vaardigheid over de verschillende leraren

methode	leraar	aantal lln.	percentage wiskundetaal-fouten		wiskundige vaardigheid	
			gemid.	*st.dev.*	*gemid.*	*st.dev.*
Getal & Ruimte	1	15	22.20	29.82	10.07	2.76
Getal & Ruimte	2	16	7.81	18.67	14.13	2.42
Getal & Ruimte	3	16	20.31	30.58	8.44	2.13
Getal & Ruimte	4	13	24.39	35.16	11.08	3.15
Getal & Ruimte	5	15	27.20	25.74	7.47	2.33
Sigma	6	29	10.90	24.50	11.17	2.78
Sigma	7	14	20.21	32.16	7.29	2.13
Moderne Wiskunde	8	34	11.47	17.71	12.35	2.73
Moderne Wiskunde	9	12	5.50	12.85	9.92	3.50
Moderne Wiskunde	10	13	12.15	19.99	10.00	3.24
Totaal		177	15.37	25.10	10.53	3.32

gen de voorspelde richting in.

De gevonden resultaten komen in grote lijn overeen met die van het in de inleiding genoemde onderzoek van De Leeuw e.a. en dat van Harskamp en Suhre. Ook bij deze onderzoeken werden geen methode-effecten gevonden. Toch volgen we Harskamp en Suhre niet in hun opvatting dat methoden er eigenlijk weinig toe doen. Ten aanzien van de rekenfouten kan het uitblijven van het voorspelde effect veroorzaakt zijn door de geringe frequentie van voorkomen over de totale groep. De betreffende opgaven lijken als meetinstrument voor dit type fouten minder geschikt. Bij de wiskundetaal-fouten is de frequentie hoger. Bij de onderzoeksopzet is echter een aantal kanttekeningen te plaatsen - we zullen dat doen aan het einde van deze paragraaf - op grond waarvan ook aan het uitblijven van een methode-effect bij de wiskundetaal-fouten geen harde conclusies verbonden kunnen worden. We vermoeden dat methoden een verschillende invloed op de leerresultaten kunnen hebben, maar op een meer complexe wijze dan volgens de in de inleiding genoemde analogie van Walker en Schaffarzick. We lichten dit toe met het oog op de wiskundetaal.

Wiskundetaal en didactiek

De leerlingen die meer op wiskundetaal getraind zijn maken niet minder wiskundetaalfouten. Er is zelfs een verschil in de tegengestelde richting, zij het niet significant. We zien in dit kleine verschil een indicatie van een prestatieprofiel, zoals dat volgens Van den Heuvel-Panhuizen (1990) zich zou moeten openbaren bij verfijnde toetsing. Dat profiel komt echter niet overeen met de voorspelling van Walker en Schaffarzick (1974). Zij stelden dat leerresultaten vooral een afspiegeling zijn van opname en nadruk van leerstof in een leerboek. Onze resultaten geven juist een indicatie ervoor dat het leerboek met de meeste nadruk op wiskundetaal er op dit punt het slechtst uitkomt. Volgens ons is niet alleen de aandacht die een bepaald onderwerp krijgt van belang, maar ook de didactiek, dat wil zeggen de manier van aandacht, het moment van aandacht[5]. We zien in onze resultaten een indicatie ervoor dat vroeg intensief gebruik van formele notaties en terminologie didactisch minder verantwoord is. Steun hiervoor is te vinden bij de binnen Nederland belangrijke stroming van het *realistisch wiskunde-onderwijs*. We behandelen opvattingen van Van Hiele, Freudenthal en Treffers. Van Hiele is een voorloper van deze stroming; Freudenthal en Treffers hebben het realistisch wiskunde-onderwijs theoretisch gefundeerd.

Van Hiele (1973) geeft aan, dat de taal van het onderwijs moet aansluiten bij die van de leerling. De leerling moet veel concrete ervaring met begrippen opdoen (bijvoorbeeld knippen en vouwen bij meetkundige figuren) alvorens de vaktaal kan worden geïntroduceerd.

Freudenthal (1978, 1983) pleit voor opbouw van de wiskunde vanuit de werkelijkheid naar analogie van de historische ontwikkeling. Het taalgebruik moet verlopen van actief-demonstratief naar feitvaststellend functioneel. Een dergelijke verloop zit in de volgende reeks uitspraken[6]:

- $3^2=9$ dus 3 is de wortel van 9; $5^2=25$ dus 5 is de wortel van 25; enzovoort;
- de wortel van een getal vind je door te zoeken naar een getal dat als je het kwadrateert weer het eerste getal oplevert;
- de wortel van een getal is een getal waarvan het kwadraat het gegeven getal is;
- $x=\sqrt{a} \leftarrow\rightarrow x^2=a$;
- worteltrekken is de inverse van kwadrateren.

Treffers (1987) beveelt (voor het aanvankelijk reken- en wiskunde-onderwijs) uitstel van verkorting aan van de notaties bij algoritmen. Algoritmen moeten zoveel mogelijk door de leeringen zelf geconstrueerd worden. Bij zowel Van Hiele als bij Freudenthal en Treffers vinden we dat een zeer actieve rol van de leerling vereist is bij het opbouwen van de eigen kennis.

Veel kenmerken van het realistisch wiskunde-onderwijs, met name wat betreft de rol van de wiskunde-taal, zien we terug in het nieuwe leerplan voor de onderbouw dat in 1993 wordt ingevoerd. We zullen de plaats van de wiskundetaal in dat leerplan kort bespreken.

De rol van wiskundetaal binnen het nieuwe leerplan voor de onderbouw

In 1993 gaat parallel aan de invoering van de basisvorming een nieuw leerplan van start voor de onderbouw van het voortgezet onderwijs: Wiskunde voor 12- tot 16-jarigen (W 12-16). Het is het sluitstuk na eerdere vernieuwingen in

het basisonderwijs en de bovenbouw van het voortgezet onderwijs. Het nieuwe leerplan staat niet alleen voor inhoudelijke vernieuwing, maar stuurt ook aan op didactische verandering, gebaseerd op het realistisch wiskunde-onderwijs. Didactische kenmerken zijn onder andere: actieve inbreng van leerlingen in het eigen leerproces met behulp van eigen constructies; progressieve mathematisering, waarbij dezelfde problemen op een telkens hoger abstractieniveau worden opgelost; ontwikkelen van modellen, schema's en symbolen; verhoging van het niveau van de leerling vanaf het niveau van de leerling zelf. De leerplanontwikkelaars zijn op het punt van het uitstellen van formele wiskundetaal erg ver gegaan. In het ontwikkelde voorbeeldmateriaal, dat als inspiratie diende voor de leerboekauteurs, is de formele wiskundetaal vrijwel geheel afwezig (zie Meijer, Kok, Riemersma & De Vries, 1991). Uiteindelijk bepalen de leerboekauteurs zelf wat er van deze didactische ideeën in hun methoden tot uitdrukking komt. Ons inziens moet enige formele wiskundetaal als kenmerkend onderdeel van de wiskunde wel behandeld worden, zij het pas aan het eind van de onderbouw. We lichten dit toe.

Bij het oplossen van problemen binnen de wiskunde speelt de wiskundetaal een belangrijke rol in de verschillende fasen van het oplossingsproces. We kunnen de volgende fasen in het oplossingsproces onderscheiden (naar Riemersma, 1991):

- encodering ('opnemen' van de opgave of taak);
- representatie-ontwikkeling (het probleem ontwikkelen, begrijpen, er vat op pogen te krijgen);
- afwikkeling (uitrekenen, algoritmisch oplossen);
- evaluatie.

Bij het encoderen moet de taal waarin het probleem gesteld is begrepen worden. Bij de representatie-ontwikkeling kunnen we onder meer denken aan het overgaan op een andere notatie. Het advies een andere notatie te gebruiken vormt één van de hints uit het door Polya ontwikkelde systeem van heuristieken (Polya, 1957). Een andere notatie kan andere aspecten van een begrip benadrukken die in een bepaalde situatie meer relevant zijn (zie ook Skemp, 1971). In de afwikkeling spelen notaties een technische rol. Bij de evaluatie moet men over een adequate terminologie beschikken om over het probleem, de oplossingsweg en de oplossing na te denken. Niet genoemd nog hierbij is het aspect van de communicatie over het probleem met anderen (de leraar of medeleerlingen). In eerder onderzoek (Perrenet, 1992) vonden we aanwijzingen, dat groepswerk een positieve invloed heeft op het correct hanteren van formele notaties.

Voor het oplossen van wiskundige problemen is het beheersen van de wiskundetaal dus onontbeerlijk. In hoeverre de weg tot de wiskundetaal in de praktijk van het wiskunde-onderwijs kan lopen via een zelf ontwikkelde wiskundige taal van de leerlingen is nog een open vraag. Meijer e.a. (1991) wijzen op het probleem van de op intuïtive noties gebaseerde vakinhoudelijke misconcepties die een belangrijke barrière in het leerproces kunnen vormen. Ons inziens is dat geen argument tegen een dergelijke onderwijsmethode; hoe eerder dergelijke concepties aan het licht komen, des te eerder het onderwijs erop kan inspelen. Een groter probleem lijkt ons dat een dergelijk leerproces slechts voor een klein deel door leerboeken gestuurd kan worden en veel (mogelijk te veel) inbreng van de leraar vereist.

Kanttekeningen bij het beschreven onderzoek naar specifieke fouten

We bespreken achtereenvolgens foutenanalyse als techniek om methode-effecten te onderzoeken, het probleem van de implementatiefactor en het effect van de leraar en andere groeperingsfactoren.

Een mogelijk kritiekpunt op de in dit artikel beschreven studie is de betrouwbaarheid van de meting bij de fouttypen. Het categoriseren van fouten in schriftelijk leerlingmateriaal heeft altijd een zeker subjectief karakter; soms lijkt het bestuderen van de uitwerkingen op protocolanalyse. Een interbeoordelaarsbetrouwbaarheid is in deze éénmansstudie niet bepaald, maar de gegeven operationalisaties vormen een objectiviteitscriterium. Daarnaast is het aantal meetpunten bij de schalen voor rekenfouten en wiskundetaal-fouten beperkt (het gaat om percentages ten opzichte van enkele opgaven). Bij de wiskundetaal-fouten berust de analyse op relatief de meeste opgaven. De grootte van de standaarddeviaties (paragraaf 3)

verplicht toch tot voorzichtigheid bij het hanteren van de resultaten. In de inleiding werd de opvatting van Gravemeijer e.a. (1990) weergegeven, dat methoden wel degelijk verschillend effect hebben, maar dat voor het meten van de verschillen verfijnde toetsen nodig zijn. Foutenanalyse gekoppeld aan het toetsen van specifieke hypothesen, zoals in deze studie, lijkt daartoe een geschikte techniek. Wel moeten dan toetsen ontwikkeld worden gericht op het meten van de desbetreffende fouten. Analyse van bestaande toetsgegevens is als basis niet sterk genoeg.

Effecten kunnen alleen bij getrouwe implementatie aan de methoden worden toegeschreven. Deze studie maakt gebruik van de data van het transfertestonderzoek (De Leeuw e.a., 1988). Als enige meting op de factor implementatie werd in dat onderzoek gevraagd aan de betrokken leraren in hoeverre men tevreden was over het gebruikte leerboek. De antwoorden differentieerden te weinig om de factor in de analyses op te nemen (vrijwel ieder was redelijk tevreden). Nederlandse wiskundeleraren volgen over het algemeen het gebruikte leerboek goed qua leerstofaanbod en de aard van de methode (zie Meijer e.a., 1988; Van Streun, 1990 en Gravemeijer e.a., 1990). Wanneer het nieuwe experimentele methoden betreft, kunnen implementatieverschillen echter wel degelijk optreden (zie bijvoorbeeld Herfs, Mertens, Perrenet & Terwel, 1991). Bij het transfertestonderzoek betrof het al langer in gebruik zijnde methoden.

Ook bij getrouwe implementatie kunnen kenmerken van de leraar los van het leerboek van invloed zijn op het leerresultaat. Daarnaast is invloed mogelijk van andere groeperingsfactoren zoals de klas en de school. We denken hierbij aan klassecompositie en schoolklimaat (zie Van den Berg, Terwel & Wierstra, 1993). Opname van meerdere groeperingsfactoren in dezelfde analyse was binnen het gegeven design van deze studie (zie Figuur 3) niet mogelijk. Er onstaat dan een tekort aan vrijheidsgraden. Binnen de Getal & Ruimte-groep kan geen onderscheid gemaakt worden tussen klas, leraar en school. Gedegener controle voor groeperingsfactoren vereist een meer gebalanceerd design.

Door de veelheid van aspecten, waarmee moet rekening worden gehouden, is het moeilijk om bij het onderzoek naar methode-effecten met overtuigende resultaten te komen. Dat blijkt ook uit de in de inleiding genoemde discussie van Harskamp en Suhre (1990) met Van den Heuvel-Panhuizen (1990). We sluiten ons aan bij hun conclusie, dat om tot deugdelijke methode-evaluatie te komen - en daarmee tot gefundeerde adviezen voor verdere ontwikkeling van methoden - samenwerking tussen didactici en onderwijskundigen zeer is aan te bevelen.

Noten

1 In De Leeuw e.a. (1988) wordt de betreffende test 'test voor actuele beheersing' genoemd.
2 Inmiddels zijn de onderzochte edities door nieuwe vervangen met deels andere kenmerken.
3 Als tegenpool van algoritmisch wordt door De Leeuw de term heuristisch gebruikt.
4 We moeten spreken van een trend vanwege de oorspronkelijke rechts-eenzijdige toetsing.
5 Walker en Schaffarzick waarschuwen zelf al tegen een al te simpele interpretatie van de door hen geformuleerde principes.
6 De uitspraken zijn als definities onvolledig: het aspect van het teken van de wortel van een getal is buiten beschouwing gelaten.

Literatuur

Berg, G. van den, Terwel, J., & Wierstra, R. F. A. (1993). Evaluatie van leergangen. In: W. J. Nijhof, H. A. M. Franssen, W. Th. J. G. Hoeben & R. G. M. Wolbert (Eds.). *Handboek curriculum: modellen, theorieën, technologieën* (pp. 371-392). Amsterdam/Lisse: Swets & Zeitlinger B.V.

CITO (1982). *Leerdoelgerichte toetsen Wiskunde deel 1*. Arnhem: CITO.

CITO (1986). *Leerdoelgerichte toetsen Wiskunde deel 2*. Arnhem: CITO.

Freudenthal, H. (1978). *Weeding and Sowing. Preface to a Science of Mathematics Education*. Dordrecht/Boston: Reidel.

Freudenthal, H. (1983). *Didactical Phenomenologie of Mathematical Structures*. Dordrecht: Reidel.

Gravemeijer, K. P. E., Heuvel-Panhuizen, M. van den, & Ploeg, D. A. van der (1990). *Ander Rekenboek Andere Rekenprestaties. Een vergelijking tussen twee reken-wiskundemethoden in groep 3. MORE-project.* Utrecht: OW&OC en VOU, Rijksuniversiteit Utrecht.

Harskamp, E. (1988). *Rekenmethoden op de proef gesteld* (Academisch proefschrift). Groningen: RION.

Harskamp, E., & Suhre, C. (1990). Kritiek op methodenonderzoek tussen wal en schip. *Pedagogische Studiën, 67,* 284-290.

Herfs, P. G. P., Mertens, E. H. M., Perrenet, J. Chr., & Terwel, J. (1991). *Leren door Samenwerken.* Amsterdam/Lisse: Swets & Zeitlinger.

Heuvel-Panhuizen, M. van den (1990). Rekenmethoden-onderzoek op de proef gesteld. *Pedagogische Studiën, 67,* 271-283.

Hiele, P. M. van (1973). *Begrip en inzicht.* Purmerend: Muusses.

Leeuw, L. de (1979). *Leren Probleemoplossen* (Academisch proefschrift). Lisse: Swets & Zeitlinger.

Leeuw, L. de, Meijer, J., Perrenet, J. Chr., & Groen, W. E. (1988). *De conctructie en validering van een transfertest voor wiskunde-onderwijs met gebruikmaking van items met gefaseerde hulp.* Amsterdam: Vrije Universiteit.

Meijer, J., Perrenet, J. Chr. & Riemersma, F. (1988). Leren probleemoplossen in het wiskundeonderwijs. *Pedagogische Studiën, 65,* 16-31.

Meijer, J. (1990). Effects of teaching methods on mathematical abilities of students in secondary education compared by means of a transfertest. In *Proceedings of the Fourteenth International Conference for the Psychology of Mathematics Education* (pp. 159-166). Oaxtépec, Mexico: Program Committee of the PME.

Meijer, J., Kok, F., Riemersma, F. S. J., & Vries, J. de (1991). *Vernieuwing van het wiskunde-onderwijs in de onderbouw van het voortgezet onderwijs.* Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam.

Meijer, J., Perrenet, J., & Groen, W. (1992). Leerboekeffecten: verkeerd weergegeven feiten en onjuiste interpretaties van Van Streun. *Euclides, 67,* 153-155.

Perrenet, J. (1992). Fouten bij het oplossen van wiskundeproblemen en de invloed van groepswerk. *Tijdschrift voor Didactiek der ß-wetenschappen, 10,* 112-125.

Polya, G. (1957). *How to solve it.* Princeton: Princeton University Press.

Riemersma, F. (1991). *Leren oplossen van wiskundige problemen in het voortgezet onderwijs* (Academisch proefschrift). Amsterdam: Stichting Centrum voor Onderwijsonderzoek, Universiteit van Amsterdam.

Skemp, R. (1971). *The Psychology of Learning Mathematics.* Aylesbury: Hazel Watson & Viney Ltd.

SPSS Inc. (1988). *SPSS/PC+ V3.0.* Chicago/Gorinchem: SPSS Inc.

Streun, A. van (1989). *Heuristisch Wiskunde-Onderwijs. Verslag van een onderwijsexperiment* (Academisch proefschrift). Groningen: Rijksuniversiteit Groningen.

Streun, A. van (1990). Leerboeken: feiten en interpretaties. *Euclides, 66,* 55-60.

Treffers, A. (1987). *Three Dimensions. A Model of Goal and Theory Description in Mathematics Instruction – The Wiskobas Project.* Dordrecht: Reidel.

Walker, D. F., & Schaffarzick, J. (1974). Comparing curricula. *Review of Educational Research, 44,* 83-111.

Manuscript aanvaard 5-3-1993

Auteur

J. Chr. Perrenet is werkzaam bij het Instituut voor de Lerarenopleiding van de Universiteit van Amsterdam.

Correspondentie-adres: Universiteit van Amsterdam, Instituut voor de Lerarenopleiding, Herengracht 256, 1016 BV Amsterdam.

Abstract

The influence of the textbook on the occurrence of specific errors in mathematical problem solving

J. Chr. Perrenet. Pedagogische Studiën, 1993, 70, 195-205.

De Leeuw et al. (1988) found no effect of any of three Dutch mathematics textbooks, with different didactical conceptions, on a mathematical transfer test. The transfer test had been presented to 325 students (age 15) with ten teachers after three years of secondary education.

This article presents the analysis of the occurrence of specific errors in the written solutions of the transfer test problems. In the analysis factors teacher and mathematical ability, as measured by an instructions problem test, have been controlled for. Again no effect was found. Students using the textbook which emphasizes notation and terminology the most and the earliest even committed notation and terminology errors more frequently, although the difference was not significant. The role of formal terminology and formal notation in mathematics education is discussed in reference to curriculum reform in The Netherlands.

Discussie

De dood van de universiteit

J. Masschelein, Katholieke Universiteit Leuven

Samenvatting

De universiteit is dood en daarmee verdwijnt ook een vormingsideaal dat zich richtte op veralgemeenbaarheid (van kennis) en op het overstijgen van het particuliere. Vandaag lijkt vorming nog slechts het particuliere belang en het zelfbehoud te dienen, en is het hoogste doel van de universiteit niet kennis, maar prestige. De universiteit is geen *instituut* van vorming meer, maar een leer*organisatie* waar communicatienetwerken de plaats innemen van het discussieseminarie. Misschien is er nog leven voor de universiteit indien ze een nieuwe vorm vindt die speelruimte laat voor een kritisch en 'storend' denken, voor verontrustende vragen en daarmee voor vorming.

1 Hoger onderwijs: zonder universiteit, zonder vorming?

De economische en instrumentele logica lijkt inmiddels het hele domein van kennis, leren en onderwijs te hebben ingenomen. De laatste fase van die invasie begon zowat honderdtwintig jaar geleden in Amerika met de introductie van 'efficiency engineers' in het hoger onderwijs, d.w.z. boekhouders die competent zijn om ondernemingen te adviseren hoe het best hun 'output' te verhogen per kosteenheid. Ze lijkt haar voltooiing te vinden met het verdwijnen, voor onze ogen, van de universiteit. De universiteit verdwijnt en dat betekent dat ook een bepaald idee van 'vorming' verdwijnt. De inzet van de discussie omtrent de universiteit (versus beroeps-hogeschool) is inderdaad niet beperkt tot het al dan niet behouden van deze instelling, maar raakt de hele sfeer van leren en onderwijzen omdat het precies gaat om een bepaald vormingsideaal. Datzelfde ideaal is ook in het spel in de discussies omtrent algemene vorming en beroepsvorming, fundamenteel en toegepast onderzoek, intellectuelen en professionelen. In feite gaat het daarbij om de plaats die de technologie, de gerichtheid op economisch nut en op particuliere belangen toekomt. Die discussie lijkt achterhaald te zijn.

De wijze waarop vandaag over de universiteit en het hoger onderwijs wordt gesproken, en de wijze waarop beide functioneren getuigen van de feitelijke overwinning van de idealen van efficiëntie, prestige, mobiliteit en professionaliteit. Dat zijn de idealen van een industriëel bedrijf of een economische organisatie en hun overwinning betekent het einde van een ideaal dat de universiteit soms nog publiekelijk belijdt: het ideaal van de speurtocht naar waarheid, van de liefde voor de kennis om de kennis. Het betekent het einde van de universiteit als institutie van geleerdheid. De universiteit bestaat niet meer, of beter de idee 'universiteit' is dood! Ze bestaat niet meer omdat ze het onderzoek zelf van elke notie van 'verplichting', gebonden aan een aanspraak, heeft ontdaan, omdat ze onderwijs en onderzoek feitelijk heeft gescheiden, en omdat ze het 'denken' heeft overgeleverd aan de logica van de communicatie en de informatie. De universiteit bestaat niet meer vanaf het moment dat het onderscheid met andere vormen van hoger onderwijs alleen nog gedefinieerd wordt in termen van graadverschil (diepgaander, fundamenteler, algemener, wetenschappelijker, op een hoger niveau), en niet in termen van verschil in 'idee', in 'bestemming', in 'essentie' of soort. Of dat erg is ? Laat mij proberen de inzet wat scherper te omlijnen. Indien daarbij overigens de schijn ontstaat van corporatistische belangenverdediging (ik werk aan een universiteit), dan is precies dàt een teken van de toestand waarin we verkeren, van het feit dat universiteit en hoger (beroeps-)onderwijs zich inmiddels op elkaars terrein bevinden.

2 De idee universiteit

De idee van universiteit leverde een normatief begrip van menselijke vorming. Die vorming betrof de handelings- en oordeelsbekwaamheid in een concrete situatie. Ze betrof de mogelijkheid 'juist', d.w.z. *menselijk* te handelen en te oordelen. 'Juist' is niet alleen technisch

juist, maar ook menselijk gerechtvaardigd, geldig. Wanneer is het handelen en oordelen geldig ? Het is geldig (menselijk) wanneer het veralgemeenbaar is, d.w.z. wanneer het teruggevoerd kan worden tot, of gefundeerd worden op veralgemeenbare principes, op beginselen die voor alle mensen kunnen gelden (die universeel zijn). Die principes en beginselen kunnen niet opgeleverd worden door de ervaring (aangezien ervaring noodzakelijk particulier en contingent is), maar zijn het werk van de 'rede', van het inzicht, zijn een kwestie van 'denken'. En, zoals Kant schrijft, het denken "geht auf die Prinzipien hinaus, aus denen alle Handlungen entspringen" (Kant, 1982, p. 707). Het meest algemene beginsel, en dus het meest menselijke beginsel, is precies het ratiobeginsel zelf, volgens hetwelke we gehouden zijn rekenschap af te leggen van ons handelen en oordelen. Het ratiobeginsel sticht de universele menselijkheid én sticht de universiteit (Derrida, 1990). De universiteit heeft niet het privilege van de menselijke vorming, maar ze is wel de enige plaats die zijn eigenheid uitsluitend dankt aan het 'ratiobeginsel'. De universiteit is (in deze welbepaalde, traditionele opvatting) vormend, omdat het funderen van het bijzondere op het algemene, of het verklaren van effecten door hun oorzaken de definitie zelf is van rationaliteit en van het mens-zijn (Hellemans, 1992, p. 8)[1].

Vorming is dus het verwerven van kennis die veralgemeenbaar is, die geen particulier belang (of nut) dient, maar op zich nutteloos of belangeloos is. De nuttigheid is een neveneffect. Vorming heeft betrekking op een veralgemeenbaar weten dat geen utilitair doel dient, maar de "formation de l'esprit" (Ariès), de algemene cultuur. Vorming als menselijke vorming kan in beginsel niet op nut gericht zijn, omdat elk nut noodzakelijk verwijst naar particuliere individuen en groepen, naar particuliere belangen, terwijl het menselijke precies datgene is wat voor alle mensen, algemeen, *geldt*. Vorming is dus gericht op de kennis als kennis, op de kennis van grondbeginselen, op de 'zuivere', 'theoretische' kennis waar elke *mens* van nature naar streeft (Aristoteles). Die kennis wordt niet verworven door overdracht of doorgeven, maar is een permanente en eeuwige opdracht, een eeuwige onderzoeksopdracht. Een opdracht die in de eerste plaats de zaak is van de 'zuivere' wetenschappen, de 'algemene' vakken, of het 'fundamenteel' onderzoek. In het opnemen van deze onderzoeksopdracht realiseert de mens zich als rationeel wezen en dus tegelijkertijd zichzelf als mens. De speurtocht naar kennis, de algemene cultuur, is daarom altijd tegelijk vorming van het menselijk subject (Lefort, 1982, p. 221). De kennis sticht menselijke identiteit, en maakt het mogelijk te oordelen en te handelen op een geldige, menselijke wijze. Ze maakt niet alleen het persoonlijke oordeel mogelijk, maar als menselijk oordeel verbindt het de persoon tegelijk met de andere mensen. De humane vorming is altijd de vorming van persoon én (menselijke) gemeenschap.

Vorming wordt dus niet algemeen genoemd omdat ze overal bruikbare of toepasbare kennis zou betreffen. Het algemene verwijst naar het gemeenschapsvormende en betreft de normatieve eis (van het ratiobeginsel) tot het rekenschap afleggen ten aanzien van allen. Dat deze vorming tegelijk ook kritisch is, betekent dat ze samengaat met een gevoeligheid voor het niet-vanzelfsprekend zijn van het concreet-gegevene en contingente, en voor de vraag van de verantwoording ervan gemeten aan het recht doen aan de 'zaak' zelf, aan wat die zaak in beginsel is. Kritiek is de toets van die rechtvaardigheid of geldigheid.

Tegen deze achtergrond wordt duidelijk waarom geen hogeschool haar universitaire pretenties kan waarmaken indien ze niet voldoet aan twee eisen: wetenschappelijk onderzoek én wetenschappelijk onderwijs, en dat in verbinding met elkaar. Het meest wezenlijke is inderdaad dat het onderwijs met wetenschappelijk onderzoek *gecombineerd* moet worden, het volstaat niet dat het erop gebaseerd is. Het gaat niet om de overdracht van wetenschappelijke inzichten. Het onderwijs is gericht op de actieve speurtocht naar kennis, daardoor is het vormend, maar relatief ineffectief voor andere doeleinden.

De 'scène' waar deze speurtocht, dit 'denken', zich afspeelde was niet de dialoog van de (Griekse) academie, maar het 'freies Diskussionsseminar' (Von Humboldt), dat gekenmerkt werd door *gedeelde* communicatieve omgangsvormen en dat geldig weten genereerde. Geldig weten is immers niet hetzelfde

als weten dat 'werkt' of nuttige effecten heeft. Het kan ook niet in besloten ruimten, laboratoria, tijdens privégesprekken, of door simpele verwijzing naar of goedkeuring van een autoriteit (een professor) 'gevormd' worden. De geldigheid wordt niet gegarandeerd door individuele competentie, maar door de mogelijkheid tot kritische discussie en de noodzaak tot verantwoording: dus door de publieke vormen van argumentatie en communicatie die op een welbepaalde plaats en binnen welbepaalde inhoudelijke tradities gevoerd worden. Die tradities zijn symbolische gehelen, 'culturen', die tegelijk de mogelijkheid bieden van geldig spreken (men kan niet om het even wat zeggen) en de basis vormen van vernieuwing. Het 'freie Diskussionsseminar', was voor Von Humboldt het model bij uitstek voor 'vorming'. Het weten wordt er niet enkel overgedragen, maar komt er tot stand, krijgt er vorm, en met die vorming worden tegelijk de menselijke persoon en de menselijke gemeenschap gevormd. Die vrije discussie is niet het privilege van de universiteit, maar de universiteit is de enige instelling waarvan de eigenheid precies aan deze scène gekoppeld is. De universiteit is een instituut gericht op de kennis als kennis en als dusdanig vormend. Ze representeert een idee en genereert *zelf*, als instituut, weten, cultuur die vormend is voor individu en gemeenschap.

3 Alles in functie van particulier belang en zelfbehoud

Het einde van de idee van universiteit vinden we o.m. gedocumenteerd in het *Memorandum on Higher Education in the European Community* dat de 'Task Force for Human Resources' heeft uitgebracht in 1991, en waarin het woord 'universiteit' nog nauwelijks voorkomt. Dit memorandum is geen uitzonderlijk document, het brengt een staalkaart van de wijze waarop vandaag binnen en buiten het hoger onderwijs over dat onderwijs wordt gedacht en hoe het functioneert. De onderwijsinstituten die het Memorandum voor ogen heeft, hebben enkel nog een instrumentele (nuttigheids-)waarde, zowel voor het individu, als voor de (Europese) gemeenschap. Het onderwijs moet de vaardigheden, technieken en kennispakketten leveren die zowel het individu, als de (lokale, nationale, Europese) gemeenschap in staat stellen zich uit te rusten en te wapenen voor het veroveren en behouden van een eigen plaats in de wereld: "defining its position on the world scene", "strengthen its world presence". Het onderwijs en de kennis dienen het particuliere belang van individuen en groepen. De onderwijsinstellingen dragen bij tot de ontwikkeling van 'human resources' en 'human capital', en zijn zelf ook een 'resource' in het realiseren van een Europese 'excellence' ter zake. Ze maken deel uit van een 'learning society' waar aanpassen, flexibiliteit cruciaal zijn en leren leren de hoogste vereiste is omdat niets (wat men leert) er echt toe doet, of tenminste er alleen maar toe doet in zoverre het een instrument is tot zelfbehoud en zelfontplooiing, tot bevestiging en versterking van de *eigen* identiteit. De gezochte identiteit en zelfontplooiing is echter volledig *afhankelijk* van de concurrentiepositie en het succes op de economische markt.

De moderne technologische maatschappij is blijkbaar een maatschappij waarin leren ter beheersing en controle centraal staat, maar waarin elk besef van het waartoe ontbreekt of niet boven het zelfbehoud uitstijgt. Alles lijkt een 'value' geworden te zijn - en, tussen haakjes, het spreken in termen van waarden, bv. menselijke waarden is geen correctie van het economische denken maar een ultieme uiting ervan. Alles verwijst naar iets anders waarvoor het geruild kan worden, een verwijzing die uiteindelijk eindigt in het pure overleven als de ultieme 'waarde'.

In dezelfde economische logica moeten individu en gemeenschap overigens een belangrijke 'resource' trachten te verzilveren: ze moeten de Europese culturele *erfenis* veilig stellen en ontwikkelen. Die 'cultural heritage' moet goed bewaard en beheerd worden, de diversiteit (in culturele deposito's en effecten) moeten behouden blijven zodat een sterke portefeuille kan ontstaan, die te gelegener tijd gebruikt kan worden ter bevestiging van de eigen positie. Cultuur wordt zelf een te verhandelen 'goed', een 'waar'.

Ook al spreekt het memorandum terloops over algemene en kritische vorming dan alleen omdat het algemene en kritische zelf elke normatieve connotatie (d.w.z. elke band met verbindende normen of met inhoudelijke tradities) verloren hebben, en louter instrumenteel zijn

gedacht. De "independent judgement, creativity and 'esprit critique'", wordt gemeten aan haar succes op de economische markt en daarmee feitelijk onderworpen aan haar imperatieven.

'Vorming' is blijkbaar zelf een goed geworden en cultuur vervallen tot een erfeniskwestie waarbij de vraag naar de geldigheid en de rechtvaardigheid zelfs niet meer opduikt. In feite gaat het memorandum over het organiseren van het particuliere zelfbehoud. Het buigt zich over middelen, strategieën, objectieven, over beschikbare, te voorziene en te ontwikkelen 'resources', over kapitaal, over transmissie- en produktieprocessen, over de vereiste (onderwijs)technici, over de gewenste produkten, over procedures voor kwaliteitscontrole, over de mogelijke cliënten, over rendement, mobiliteit, effectiviteit en efficiëntie, over 'vitale' kwesties...

Het gaat overigens niet meer om onderwijs*instituten*, maar om leer*organisaties*[2], organisaties waar bepaalde capaciteiten tot handelen en oordelen doorgegeven en geproduceerd worden, bedrijven dus, of ondernemingen. De referentie is dan ook niet de geldigheid in de context van de cultuur, maar de prestatie in de context van de markt. De betekenis van de cultuur wordt hier niet begrepen als iets dat het Subject letterlijk vormt, maar eerder als een goed dat verworven kan worden en dat verzilverbaar is in de strijd om het zelfbehoud. De organisatie neemt wel cultuur op of geeft ze door, maar als 'meerwaarde' en dus niet in de strikte zin als identiteits*vormend*. De nadruk komt te liggen op economische produktiviteit en op de celebratie en verdediging van het belang en de voortreffelijkheid of excellentie van zichzelf, zijn groep, zijn gemeenschap, zijn universiteit.

4 Prestige als hoogste doel

Bij het lezen van het *Memorandum on Higher Education in the European Community* is het moeilijk niet te denken aan dat andere Memorandum, het *Memorandum on the Conduct of Universities by Business Men* dat door Veblen in 1918 onder de titel: *The Higher Learning in America* werd gepubliceerd. Ik had dat hier eigenlijk graag integraal geciteerd, maar ik beperk mij tot Veblens argument dat de overheersing van het ondernemingsideaal in de wereld van kennis en leren rampzalig is voor de speurtocht naar waarheid, voor het wetenschappelijk onderwijs-en-onderzoek, kortom voor de universiteit. De overheersing komt feitelijk tot uiting in de centrale rol die de academische overheid speelt en in het allesbepalende belang van de 'faam' van de instelling. Faam of prestige is het competitieve doel dat de universiteit of de instelling van hoger onderwijs nastreeft. Het is op dat prestige dat het management van deze 'onderneming' vooral gericht is, en het is daartoe dat het *beginsel van standaardisatie en controlerende boekhouding* wezenlijk is.

Nu heeft deze bekendheid als dusdanig geen enkele substantiële betekenis voor het werk van de universiteit als universiteit. Op zich genomen bevordert het de speurtocht naar kennis of het denken helemaal niet. Maar men neemt gewoonlijk aan dat bekendheid als consequentie heeft dat er, via een toename van middelen en van uitwisseling, meer mogelijkheden geschapen worden voor het eigenlijke werk van de universiteit: wetenschappelijk onderzoek-en-onderwijs. Vandaar dat al wat niet dit doel van prestige dient, geen voet aan de grond krijgt in het heersende academisch beleid, en omgekeerd alleen dat wat publicitaire waarde heeft en tot het prestige bijdraagt een kans krijgt. De carrière van de leden en van de groepen binnen de onderneming is daaraan gebonden. Die leden zijn inmiddels overigens bedienden geworden met hun eigen corporatieve belangen. De stimulering en tegelijkertijd de controle en bewaking van de 'publieke' optredens van haar leden, en in het algemeen van dat wat haar faam bezorgt, is de grootste plicht van de onderneming. De principes en standaarden van organisatie, controle en prestatie die geaccepteerd worden als normaal in de bedrijfswereld, worden nu ook onvermijdelijk en noodzakelijk geacht met betrekking tot het domein van leren en onderwijs. De plichten van de beleidsman zijn enerzijds publiciteit en zijn grootste kwaliteiten zijn hier "statistical display, spectacular stage properties, vainglorious make-believe and obsequious concessions to worldly wisdom" (Veblen, p. 240). Anderzijds heeft de beleidsman vooral administratieve plichten die te maken hebben met de coördinatie en het af-

stemmen van de onderdelen van de organisatie, met disciplinering en met controle en standaardisatie van het personeel en hun taak. Deze plichten zijn er alleen maar omdat er een opdracht is tot centrale controle, zonder die noodzaak zou dat helemaal niet hoeven. Wanneer geen centraal management nodig is, dan kunnen heel wat mensen en vooral taken verdwijnen.

Men redeneert dat het prestige betekent meer fondsen én meer studenten (en dus ook weer meer fondsen). Men gelooft dus dat men zonder deze reputatie geen fondsen en geen studenten zal kunnen verwerven, en dat dat een verlies van mogelijkheden zou betekenen voor het wetenschappelijk onderzoek-en-onderwijs.

Welnu, men mag aannemen dat, zoals bij andere ondernemingen, de 'faam' wel een element is van competitie, maar geen aggregatie-effect heeft, d.w.z. dat ze geen supplementaire middelen mobiliseert voor het domein waarop de onderneming werkzaam is. M.a.w. voor de zaak van het onderwijs en het leren in het algemeen, en daarom ook voor de menselijke gemeenschap voor zover ze geïnteresseerd is in de vooruitgang van de kennis, komt er geen waarneembare nettowinst voort uit deze competitieve publiciteit. Misschien dat er een zeker effect is van het wijzen op het (economisch of het zelfbehoud dienende) belang van wetenschappelijk onderzoek-en-onderwijs, maar de toename aan fondsen is niet evenredig met de toename aan investeringen (personen, tijd, diensten, logistiek) in het verwerven van 'naambekendheid'. Als geheel genomen kosten die meer dan ze opbrengen. Het belangrijkste effect lijkt te zijn dat de stroom van middelen zich verlegt van de ene instelling naar de andere, of binnen de universiteit van de éne (vak)groep naar de andere (op deze wijze wordt het belang van 'faam' overigens verdedigd). Bovendien moet men om te kunnen wedijveren, vergelijkbaarheid nastreven, dus standaardiseren (indexen), en ook vergelijkbare middelen inzetten: *dezelfde* apparatuur, *hetzelfde* soort personeel, *dezelfde* uitrusting (het argument voor het samengaan van instellingen is trouwens precies dat men de concurrentie zal aankunnen indien men over de middelen beschikt die vergelijkbaar zijn met andere (buitenlandse) instellingen). Opnieuw is dat alleen maar nodig voor de competitie, aan het eigenlijke werk van de universiteit draagt het niets bij.

Ook de poging om zoveel mogelijk studenten aan te trekken en het stimuleren van de uitwisseling van studenten brengt geen enkele wetenschappelijke vooruitgang op, zeker niet als het gaat om uitwisselingen met equivalente, vergelijkbare instellingen, en als men daarom de uitwisseling eigenlijk verdedigt vanuit de extra-curriculaire attractiviteit of het grote prestige van die instelling. Het brengt alleen maar een overbelasting mee van het 'corps', en een versterking van de neiging om onderwijs en onderzoek van elkaar te scheiden, en om het onderwijs in min of meer vastgelegde (en daardoor vergelijkbare) curricula te vertalen, terwijl het samengaan van onderwijs-en-onderzoek (de vorming van de geleerdheid) eigenlijk geen vastgelegd curriculum verdraagt.

In de gewone ondernemingen is publiciteit een middel tot competitieve verkoop en gerechtvaardigd door de toegenomen voordelen die de succesvolle adverteerder van het toegenomen verkeer verkrijgt, en om dezelfde redenen zijn standaardiseringen, vergelijkbaarheid, uitwisselbaarheid, flexibiliteit, mobiliteit enz. nodig. In de wetenschap en het domein van de kennis echter, en los van de persoonlijke ambities van de beleidsvoerders, is er niets dat overeenkomt met deze toegenomen uitwisseling of deze competitieve voordelen.

We zitten nu met universiteiten die als ondernemingen en beroepsscholen en als instellingen voor de verspreiding van nuttig advies ongetwijfeld het beste bieden, meer in het bijzonder als we hun eigen cijfers mogen geloven. Alleen is de zaak van de universiteit zelf van ondergeschikt belang geworden. En dat lijkt onherstelbaar omdat die universiteiten geëngageerd zijn in zo'n massa verplichtingen en verbintenissen, vreemd aan hun eigenlijk enige legitieme belang, dat ván de kennis, dat men zich zelfs niet meer kan voorstellen dat ze er nog iets mee te maken hebben. Haar leden zijn inmiddels al lang geen intellectuelen meer maar prof(essionelen).

5 De verandering van de scène (organisatie en netwerk)

De universiteit werd gesticht door het ratiobeginsel, de verplichting de *ratio* op te leveren, de eis het particuliere tot veralgemeenbare beginselen te herleiden en recht te doen aan de zaak zelf. Die beginselen vormden de algemene menselijke cultuur, het geldige en beschikbare weten dat als symbolisch geheel, als inhoudelijke traditie (de wereld van) het denken begrensde en menselijk handelen en oordelen mogelijk maakte. Die universiteit is dood. Het ratiobeginsel is getransformeerd tot het beginsel van standaardisatie en controlerende boekhouding (verantwoording van de *prestatie*, 'accountability'), tot de verplichting beschikbare *informatie* op te leveren, de eis de eigen activiteiten in relatie te brengen tot vergelijkbare standaarden en de eis recht te doen aan de onderneming als onderneming d.w.z. aan haar prestige. De vergelijkbare standaarden vormen het geheel van harde criteria, de 'werkelijkheid', de feiten die de 'echte' wereld uitmaken, die buiten (de wereld van) het denken valt, en het enige criterium biedt voor de betekenis van het handelen en oordelen. Er treedt dus een verschuiving op in het criterium. Niet meer de geldigheid, niet meer de grenzen of normen van een symbolisch systeem, maar de effectiviteit, 'het lukken of mislukken' aan de feitelijkheid, aan tastbare en meetbare grenzen wordt de maatstaf voor het menselijke handelen en oordelen. De grens is niet meer het ondenkbare, maar het ondoenbare. Niet de symbolische grens, maar het feitelijke mislukken (van het experiment, van de operatie, ...), de feitelijke dood (van de instelling, van het onderzoek, van het vakgebied) is de limiet[3].

De verschuiving van het criterium tekent zich vooreerst af in de feitelijke scheiding van onderwijs en onderzoek. Dat gebeurt niet alleen door het oprichten van onderzoeksinstituten die zich buiten de universiteit gaan bewegen, maar ook door binnen de universiteit deze 'taken' te scheiden. Bovendien wordt de financiering van het onderzoek onttrokken aan de instanties (afdelingen) die ook instaan voor het onderwijs. Het onderzoek wordt onttrokken aan de normale werkingskredieten zodat de 'controle' elders gebeurt. Alle ondermaatse activiteiten moeten immers geweerd worden, en de maat is het "Centre of Excellence". De mogelijkheid van het ongelijk, van het niet-geldigzijn, is eigenlijk geen centrale bekommernis meer, maar wel het lukken of mislukken, de voorsprong of de achterstand, de excellence. Deze 'excellentie' staat, zoals het in het informatieblad van de Dienst Onderzoekscoördinatie van de K.U. Leuven vermeld wordt, boven alle 'verdenking' (GeDocumenteerd, 1992/18). Ze staat inderdaad buiten (de wereld van) het denken. Excellence betekent immers niet beargumenteerbaarheid en zinvolheid, maar wel het niet mislukken aan harde criteria die omwille van de vergelijkbaarheid en de uitwisselbaarheid en de verwisselbaarheid zo volkomen formeel en gestandaardiseerd moeten worden dat ze het karakter krijgen van 'feiten'. Excellence vereist karaktersterkte, zelfvertrouwen, doorzettingsvermogen, professionalisering, efficiëntie, mobiliteit, alertheid, ... (geleerdheid ?).

Wat is de scène waarop het wetenschappelijk onderzoek zich vandaag afspeelt? Het is niet langer de dialoog van de academie, het publieke gesprek van Humboldts "freies Diskussionsseminar", maar de uitwisseling binnen de informatie- en de communicatienetwerken. Deze netwerken creëren een zeer toegankelijke sociale ruimte, maar die ruimte wordt steeds privaat gebruikt. Ze bestaat uit een stelsel van gangen en poorten, maar niet uit een 'scène' (Verschaffel, 1991). Er zijn oneindig veel contacten, er is oneindig veel uitwisseling, maar het netwerk verbindt ze niet tot een 'plaats'. De 'scène' valt samen met het geheel van mogelijke verbindingen, maar is niet representeerbaar en daarom ook niet publiek. Dit soort communicatie tendeert er naar alle 'storingen' die te maken hebben met lokale, inhoudelijke tradities, met vragen van verantwoording tegen de achtergrond van die tradities, met het verschil tussen de plaatsen op te heffen. Niet de geldigheid is de norm, maar de totale beschikbaarheid en toegankelijkheid. Het verschil tussen de plaatsen verdwijnt, en daarmee de mogelijkheid tot kritiek in termen van argumenten en geldigheid. Er dreigt geen monopolie van de kennis, maar een gebrek aan mogelijkheden om die kennis in vraag te stellen, een gebrek aan vorming.

6 De universiteit is dood, leve de universiteit

Ik wil geen pleidooi houden voor de oude idee van universiteit, ik heb er aan herinnerd om er op te wijzen hoe dood ze wel is en hoe vals het klinkt wanneer men zich vandaag nog op haar beroept. Men mag overigens niet vergeten dat de restauratie van die idee zich vooral zou laten gebruiken in de competitie om de identiteit, om prestige. Ik wil geen pleidooi houden voor de oude idee, niet alleen omdat ze als ideaal van vorming blind is geweest voor het feit dat ze feitelijk toch particuliere belangen, belangen van elites heeft gediend, maar vooral omdat het vermoeden bestaat dat die idee wellicht in 'wezen' technisch is, "dat het ratiobeginsel niet langer (te) scheiden is van alleen al het idee van technologie" (Derrida, 1990, p. 225) en daarmee zelf mee aan de basis ligt van de overheersing van het economisch-technologische denken.

Ik wil daarom een pleidooi houden voor een universiteit en een vorming die blijvend het denken en het leren denken centraal stellen, maar dit denken niet identificeren met de speurtocht naar kennis om de kennis, naar algemene, funderende beginselen of oorzaken. Geen 'kritisch' maar een 'verontrustend' of 'storend' denken dat enerzijds wel het ratiobeginsel gehoorzaamt, maar anderzijds het funderingsdenken zelf ook in vraag stelt. Zoals Derrida opmerkt sticht het ratiobeginsel de (klassieke) universiteit, maar is de verontrustende vraag eigenlijk de vraag of deze stichting zelf gefundeerd is op dat beginsel? Deze vraag is een vraag naar de verhouding tussen het fundament en dat wat de vraag naar het fundament doet rijzen, maar zelf geen fundament is. Het denken houdt geen halt bij het fundament. Het antwoord op de verantwoording die het denken is, is hier van een principieel andere aard dan het funderen op een beginsel, ze heeft te maken met de opdracht van het openhouden van de ruimte van (een verhouding tot) het 'mogelijke', het 'andere'.

Het onderwijs mag zich niet (volkomen) laten in beslagnemen door de dringendste problemen van de samenleving (racisme, aids, drugs, milieu) of de hoogste noden (ingenieurs, verpleegsters, ...) van de arbeidsmarkt. Doelgerichtheid, nut, overleving zijn in de 'werkelijkheid' van het hoogste belang. Het onderwijs mag zich echter niet volledig door de categorie 'werkelijkheid' (de noodzaak van het individuele en collectieve overleven) laten innemen, maar moet zich ook door de categorie 'mogelijkheid' (vrijheid) laten bepalen, zich afgrenzen van de 'werkelijkheid', een 'speelruimte' laten voor reflectie en voor reflectie over de voorwaarden van die reflectie. Zijn opdracht en a fortiori die van de universiteit is dan ook niet het 'leren leren', maar het 'leren denken'. Denken dat is zich vragen stellen over wat we met enige geldigheidsaanspraak weten en kunnen, en is daarmee gebonden aan een traditie, maar denken is ook de vraag stellen naar de voorwaarden van die traditie, naar datgene wat die traditie niet is, wat er niet toe behoort. Denken dat betekent vandaag: storen.

Noten

1 Op dezelfde wijze zegt Habermas dat de universiteit die gefundeerd is op een veralgemeenbare communicatieve procedure vormend is omdat dat de definitie zelf is van de (communicatieve) rationaliteit die het wezen is van de mens (als taalwezen). Vgl. Habermas, 1986; Hellemans, 1992, pp. 8-9.

2 Het is opvallend hoe in veel literatuur over het onderwijs het onderscheid tussen instituut en (sociale) organisatie eigenlijk niet (meer) gemaakt wordt. Dat de verschuiving van instituut naar organisatie niet onschuldig is, kan men vermoeden wanneer men zich afvraagt wat het betekent om bv. het huwelijk geen instituut meer te noemen maar een organisatie in functie van particuliere behoeften of belangen.

3 De ruimte ontbreekt hier om aan te geven dat de 'klassieke' interpretatie van het ratiobeginsel als 'vraag naar de fundering', zelf deze vernietiging van de symbolische grenzen heeft bewerkt.

Literatuur

Derrida, J. (1990). De universiteit in de ogen van haar pupillen. *Comenius, 38*, 217-135.

Habermas, J. (1986). Die Idee der Universität. *Zeitschrift für Pädagogik, 32*, 703-718.

Hellemans, M. (1992, Mei). *Het wetende niet-weten.* Paper presented at the Lectures "Heeft de school nog een vormingsproject", Leuven.

Kant, I. (1982). Ueber Pädagogik. In *Schriften zur Anthropologie, Geschichtsphilosophie, Politik und Pädagogik 2.* XII, Frankfurt am Main: Suhrkamp.

Lefort, Cl. (1992). *Ecrire. A l'épreuve du politique.* Paris: Calmann-Lévy.

Task Force Human Resources (Ed.) (1991). *Memorandum on Higher Education in the European Community.* (COM (91) 349 final). Brussel.

Veblen, Th. (1965). *The Higher Learning in America. Memorandum on the Conduct of Universities by Business Men.* New York: Sentry Press, (Orig. Viking Press, 1918).

Verschaffel, B. (1990). De kring en het netwerk. Over het statuut van de publieke ruimte. In J. Masschelein (Ed.), *Opvoeding in de stad.* (pp. 47-58). Leuven: ACCO.

Manuscript aanvaard 1-3-1993

Auteur

Dr. J. Masschelein is docent en medewerker aan het Centrum voor Fundamentele Pedagogiek, K.U. Leuven.

Correspondentie-adres: Vesaliusstraat 2, 3000 Leuven

Abstract

The death of the university

J. Masschelein. Pedagogische Studiën, 1993, 70, 206-213.

The university is dead, and this entails also the disappearance of an ideal of formation aimed at universalizability (of knowledge) and at exceeding the particular. Today formation appears to serve only the particular interest and self-preservation, and the highest goal of the university is not knowledge but prestige. The university is no longer an *institution* of formation but a learning *organization*, where communication networks have taken the place of discussion seminars. There may still be life for the university, if it finds itself a new form which leaves room for critical and 'disturbing' thinking, for perturbing questions and thus for formation.

Discussie

'Storen' – De laatste strohalm van het idealisme

Reactie op J. Masschelein, 'De dood van de universiteit'

P. Baggen en I. Weijers, Katholieke Universiteit Nijmegen

Samenvatting

In onze reactie op het artikel van J. Masschelein gaan we op twee punten in zijn verhaal in. Ten eerste menen wij dat Masschelein voorbij gaat aan het fundamentele probleem van de klassieke Bildungsuniversiteit betreffende de dominantie van de filosofie. Ten tweede stellen wij voor het begrip reflectie niet naar de marge van de universiteit te verbannen maar haar in de zin van reflectie op de vakkennis juist in het centrum van haar aandacht te plaatsen.

"Die Zukunft unserer Universitäten beruht auf der Wiedererneuerung ihres ursprünglichen Geistes." Zo begon *Die Idee der Universität*, de rede van Karl Jaspers ter gelegenheid van de heropening van de Universiteit van Heidelberg in mei 1945.[1] Bijna een halve eeuw later concludeert Masschelein dat de wedergeboorte van de idee is mislukt: "De universiteit bestaat niet meer, of beter de idee 'universiteit' is dood!" Masschelein stelt de problematiek van het hoger onderwijs in hetzelfde vocabulaire als Jaspers, dat wil zeggen in de termen van het Duitse idealisme, waarbij de universiteit als gedaante van de Idee 'universiteit' wordt beschouwd: zodra de Idee uit haar ontsnapt geeft zij de geest.

De klacht van Masschelein is natuurlijk niet nieuw. Zij past in een geheel van kritische geluiden over de gang van zaken aan de universiteit, zoals de permanente ingrijpende bemoeienissen van de staat, de toenemende rol van de externe financiering van het onderzoek, de voortgaande scheiding tussen onderwijs en onderzoek en de opgedrongen concurrentie met het hoger beroepsonderwijs. Zo sprak de collegevoorzitter van de Universiteit van Amsterdam onlangs van de "verweesde universiteit": de universiteit verkeert in verwarring over haar positie en missie. Gevers ziet de toekomst echter niet zo donker als Masschelein. Volgens hem gaat het erom dat de universiteit in de komende jaren haar identiteit hervindt, dat wil zeggen de identiteit van de klassieke Europese universiteit.[2] Maar Masschelein meent juist dat een dergelijk beroep op de oude idee vals klinkt. Hij spreekt zich ten slotte uit voor een universiteit die een storend denken centraal stelt. In deze korte reactie willen we ingaan op twee punten, Masscheleins bezwaren tegen de oude idee van de universiteit en zijn alternatief van een storend denken, omdat het eigene van zijn bijdrage aan de discussie over de universiteit precies daarin lijkt te zitten.

Het probleem van de Bildungsuniversiteit

Masschelein brengt twee bezwaren naar voren tegen de oude idee van de universiteit: als vormingsideaal is ze blind geweest voor het feit dat ze particuliere belangen heeft gediend, terwijl ze bovendien zelf verbonden lijkt met 'het economisch-technologische denken'. Met deze kritiek op de klassieke Bildungsgedachte continueert Masschelein een typisch idealistische denktrant: de oorspronkelijke Idee was niet zuiver, sociaal gezien niet en principieel niet. Die kritiek komt overigens niet van harte. Er is een opvallende discrepantie tussen de voorafgaande uitvoerige en stellige litanie over de dood van de idee van de universiteit en de zeer korte en wijfelende mededeling dat er met die idee zelf ook iets niet klopte, dat "het vermoeden bestaat", "dat die idee wellicht"...

De kritiek op het klassieke Bildungsconcept had wat ons betreft wel wat royaler en meer ter zake gemogen. We willen hier in elk geval wijzen op een cruciaal probleem van dit concept dat Masschelein buiten beschouwing laat. Dit betreft de onaantastbaar en centraal geachte positie van de filosofie. De universiteit van Von Humboldt en Schleiermacher werd bijeengehouden door de filosofie: zij werd geacht zowel de eenheid der wetenschappen als de eenheid van onderwijs en onderzoek te garanderen en

zij stond tenslotte borg voor de eenheid van wetenschap en emancipatie. Aan de filosofie kwam de heerschappij toe binnen de universiteit omdat, zoals Schleiermacher het uitdrukte, "alle leden van de universiteit, tot welke faculteit zij ook mogen behoren, in haar hun wortels moeten hebben".[3]

Van meet af aan stond deze claim op gespannen voet met de ontwikkeling van de wetenschappen. Opvallend genoeg handhaaft Masschelein in zijn weergave van de lotgevallen van de universiteit zonder reserve de klassieke voorstelling van zaken. Voortdurend plaatst hij instrumentele logica tegenover vorming, en technologie tegenover "kennis om de kennis", geheel in de traditioneel Aristotelische denkstijl van een gegeven hierarchische verhouding tussen hogere, zuivere en lagere, technische kennis, epistèmè en tèchnè. In de loop van de negentiende eeuw wordt de kloof tussen epistèmè en tèchnè echter in toenemende mate gedicht. Wetenschappelijke kennis krijgt steeds sterker een proefondervindelijk karakter en wordt betrokken op uiteenlopende praktische processen en daarmee wordt het traditionele standenverschil tussen beide soorten kennis steeds moeilijker te verantwoorden.[4] De proefondervindelijke wending van de wetenschap waarbij de filosofie haar integrerende kracht en monopoliepositie wat betreft de interpretatie van de cultuur verliest wordt door Masschelein benoemd in dramatische termen van "invasie" en "dood". Wij geven de voorkeur aan een nuchterder historische observatie zoals Habermas bijvoorbeeld in dit geval de "emancipatie der empirische wetenschappen" vaststelt, waarmee we tegelijkertijd het fundamentele probleem van de Bildungsgedachte helder voor ogen houden.[5]

Een speelruimte voor reflectie?

Masschelein pleit tenslotte voor een universiteit die zich niet alleen laat bepalen door doelgerichtheid en nut, maar die ook ruimte laat voor "leren denken" of "storend denken". Met dit pleidooi lijkt Masschelein de lijn van zijn verhaal consequent door te trekken: de oorspronkelijke idee van het leren denken van het oude Bildungsideaal was niet geheel zuiver en de enige mogelijkheid lijkt nu nog slechts een vrije speelruimte te claimen, waar ongestoord en belangeloos fundamentele vragen kunnen worden gesteld. We willen over dit voorstel twee opmerkingen maken. De eerste betreft de duidelijkheid: we signaleren een aantal vage of zelfs contradictoire elementen in dit voorstel. Met name de tegenstelling tussen leren leren en leren denken blijft onhelder. Mondt leren denken niet onontkoombaar uit in leren leren aangezien denken zich per definitie op geen enkele inhoud kan vastleggen?[6] Ook de combinatie van de termen speelruimte en storen lijkt ons niet erg gelukkig. Wie of wat wordt gestoord door wat zich binnen deze "afgrenzing van de werkelijkheid" voordoet? Veronderstelt de notie storend of verontrustend denken overigens niet zelf een doelgerichtheid? En als dat niet zo is, wat betekent storen dan eigenlijk? Keert hier niet gewoon de oude onduidelijkheid van de Bildungsgedachte terug, namelijk de aanname dat politieke onthouding zou moeten samengaan met emancipatie?

Ten slotte menen wij dat het begrip reflectie in verband met de universiteit een degelijker invulling verdient dan hier wordt gegeven. Wij zouden er juist voor willen pleiten voor reflectie geen aparte, van de werkelijkheid afgegrensde speelruimte te creëren, maar daaraan een plaats midden in de werkelijkheid en midden in de universiteit toe te kennen. We denken aan reflectie op de studie en beoefening van de verschillende vakwetenschappen zelf, zowel op de status van de eigen vakkennis als op de consequenties van de maatschappelijke toepassing van die vakkennis. Reflectie niet als domein van de filosofie maar als aspect van de wetenschappelijke opleiding zoals die zich heeft geëmancipeerd tot eigen discipline. De revitalisering van de universiteit betekent niet zozeer het hervinden van haar oude identiteit of de wedergeboorte van haar idee als wel het vinden van een nieuwe bij deze tijd passende identiteit.[7]

Noten

1 K. Jaspers, *Die Idee der Universität*, oorspr. Berlijn 1946, reprint Heidelberg, New York 1980, p. 5.

2 J. K. M. Gevers, 'De verweesde universiteit'. *Krisis* 49, 1992, p. 56-61.

3 F. Schleiermacher, 'Gelegentliche Gedanken über Universitäten im Deutschen Sinn'. In E. Anrich (Ed.), *Die Idee der deutschen Universität*, Darmstadt 1959, p. 259.

4 Interessante aspecten van dit veranderingsproces worden beschreven in J. C. M. Wachelder, *Universiteit tussen vorming en opleiding. De modernisering van de Nederlandse universiteiten in de negentiende eeuw*, Hilversum 1992.

5 J. Habermas, 'De idee van de universiteit – leerprocessen'. *Comenius* 38, 1990, p. 177.

6 Schnaedelbach heeft in dit verband gewezen op het moderne uitgangspunt van Von Humboldt, dat de universiteit de wetenschap altijd als een nog niet helemaal opgelost probleem moet behandelen en daarom altijd met onderzoek verbonden blijft. "Nur das Forschen selbst hat sie als das Lehrbare zubegreifen, nicht, 'die' Wahrheit; darin besteht die 'Einheit von Forschung und Lehre'". H. Schnaedelbach, *Philosophie in Deutschland 1831-1933*, Frankfurt a.M., 1983, p. 41.

7 Een interessante eigentijdse uitwerking van de gedachte dat reflectie een centrale plaats verdient binnen het universitaire onderwijs en onderzoek biedt R. Barnett, *The Idea of Higher Education*, Buckingham, 1990.

Manuscript aanvaard 19-4-1993

Auteurs

P. Baggen en I. Weijers zijn respectievelijk als junior- en senioronderzoeker verbonden aan het Instituut voor Wijsgerige en Historische Pedagogiek van de Katholieke Universiteit Nijmegen.

Adres: Postbus 9103, 6500 HD Nijmegen

Abstract

'Interference' – Last straw of idealism Response to J. Masschelein, 'The death of the university'

P. Baggen & I. Weijers, Pedagogische Studiën, 1993, 1970, 214-216.

In our response to the article of J. Masschelein we enter into two issues. First, we think that Masschelein passes over the fundamental problem of the classical Bildungsuniversity. Here the dominance of philosophy is at stake. Secondly, we propose not to banish the concept of reflection to the margins of the university but instead to put reflection on professional knowledge in the centre of its attention.

Discussie

Het belang van het onderscheid

Reactie op P. Baggen en I. Weijers

J. Masschelein*, Katholieke Universiteit Leuven

Uiteraard hebben Weijers en Baggen gelijk. En ik geef graag toe dat bepaalde elementen in mijn betoog te weinig uitgewerkt zijn, al heeft dat ook daarmee te maken dat mijn bijdrage in de eerste plaats een reactie wou zijn op het Memorandum on Higher Education van de Europese Gemeenschap. Ik wil het gelijk van Weijers en Baggen helemaal niet betwisten, maar van de gelegenheid gebruik maken om wat meer duidelijkheid in de discussie te brengen. Ik vermoed immers dat het gelijk van Weijers en Baggen het 'hoger onderwijs' betreft (als sociologische categorie), waar ik het wou hebben over de 'universiteit'. Dat is, of beter was iets helemaal anders, en over dat andere gaat het. Het gaat over het belang van het maken van onderscheid, over aandacht voor wat in de traditionele metafysica de 'alteritas', de 'andersheid' wordt genoemd en daar een zekere transcendentie uitdrukt, maar in de socio- en psychologische wetenschappen nog slechts een kwestie van conventie en daarmee van definitie lijkt te zijn. Er zijn minstens twee manieren waarop het onderscheid tussen fenomenen (hun 'andersheid') in die 'wetenschappen' eigenlijk wordt geneutraliseerd, en bij Baggen en Weijers (en bij Habermas, waar beiden zo gretig naar verwijzen) treffen we ze allebei aan. De eerste is de historisering van fenomenen die berust op een opvatting van de geschiedenis als een proces, als een stroom waarin alles wat in die stroom is opgenomen kan veranderen in iets anders: bv. een institutie in een organisatie. De tweede is de functionalisering waarbij alles wat dezelfde functie(s) heeft ook hetzelfde is (zodat de functie het fenomeen bepaalt en niet omgekeerd): bv. alles wat de functie(s) van een universiteit heeft is ook een universiteit. Ik wil (met H. Arendt) blijven vasthouden aan het belang van het maken van onderscheid. Ik verwijs naar enkele elementen die ons minstens kunnen laten vermoeden dat het maken van een onderscheid bv. tussen 'instituut' en 'organisatie', en tussen 'kennen/weten' en 'denken' in de context van de discussie omtrent de universiteit van belang is.

* Ik heb mij in mijn reactie sterk laten inspireren door het denken van H. Arendt (o.m. in haar studies over gezag en over het leven van de geest).

1 Instituut en organisatie

Het onderscheid tussen instituut en organisatie wordt in de sociaal-wetenschappelijke literatuur nauwelijks nog gemaakt, de termen worden vaak stilzwijgend door elkaar gebruikt (bv. door Habermas in zijn tekst over de universiteit). Dat is, naar ik meen, onterecht.

Een instituut (re)presenteert een verbinding tussen mensen. Dat 'verbindende' werd in een bepaalde traditie inderdaad gedacht als een 'idee', een 'idée directrice' (Hauriou): een 'idee' van het werk of de onderneming die in een sociaal milieu verwerkelijking en rechtsbestand vindt. Die 'idée directrice' mag echter absoluut niet verward worden met het doel of de functie van de onderneming, ze betreft daarentegen de 'zin' ervan. En die 'zin' hoeft helemaal niet *in* de hoofden, of de voorstellingen van de leden aanwezig te zijn (dat denken alleen zij die, zoals Habermas, eigenlijk geen 'tussen' of geen 'verhouding' kennen). Integendeel, de 'zin' is zelfs helemaal niet voorstel-baar, ze roept de leden van het instituut en hun voorstellingen juist in het leven door die leden in een verhouding te plaatsen. Het zijn niet de formele regels en procedures, niet de functies, noch de subjecten die de basis vormen van het instituut, maar het is omgekeerd het instituut dat de basis vormt. Een belangrijk punt daarbij is precies dat het instituut 'geboren' wordt, dat het een stichting is (een gebeurtenis), een begin dat niet herleidbaar is tot wat er aan voorafgaat, maar zelf iets nieuws instelt.

Een organisatie daarentegen duidt op de geregelde, geformaliseerde realisering van individuele of collectieve doelen. Organisaties ko-

men tegemoet aan behoeften van individuele en collectieve subjecten die vooraf bestaan, de organisatie is er als het ware de reactie op (wat niet wil zeggen dat organisaties niet creatief of innovatief kunnen zijn). Dat wat een organisatie bindt is het gemeenschappelijke belang of doel. De organisatie is de geregelde omgang met feitelijke, waarneembare en vaststelbare *problemen* en niet met zin.

'Zin' is niet voorstelbaar, theoretiseerbaar, niet vast te leggen in een 'identiteit', in een begrip, een doel (en in die betekenis ook niet idealiseerbaar), maar precies datgene wat daaraan 'zin' verleent, wat ons de identiteit of het doel, letterlijk, 'zeggen' of te zeggen hebben. Ik denk dat het vandaag in het spreken over de universiteit dus (ook) zou moeten gaan over de 'zin' van de universiteit: wat 'zegt' ze ons (nog). De andersheid van het instituut 'universiteit' heeft daarmee te maken. En die andersheid moeten we nu inderdaad proberen te denken 'buiten de idee', zonder daarom te vervallen in een organisatorisch of functionalistisch denken. De 'zin' van de universiteit verwijst niet naar haar doel*stelling*, maar eerder naar haar op*gave*. En die is (of was) niet alleen te weten of te kennen, maar ook te denken.

2 Kennen/weten en denken

De opgave van de universiteit betreft (vandaag) niet enkel het weten (de wetenschap), maar ook het denken, dat, naar Kants onderscheid van rede en verstand, te maken heeft met wat niet gekend, niet geweten kan worden, wat te maken heeft met 'zin'. Het weten/kennen heeft met waarheid te maken, met dat wat dus kenbaar is, en op één of andere wijze identificeerbaar, voorstelbaar en waarneembaar. Het denken heeft met 'zin' te maken, met dat wat niet kenbaar is, niet te vatten is, niet in een identiteit, in een weten kan vastgelegd worden. Het denken levert geen kennis op. Het houdt zich niet bezig met de vraag wat iets is, of waartoe en waarom het bestaat (cfr. het ratiobeginsel waar ik in mijn tekst naar verwezen heb), maar met de vraag 'wat het betekent dat het bestaat', wat ons het bestaan ervan 'zegt'. Het gaat hier om een vraag die niet beantwoordbaar is, ook niet door het denken, maar die het denken beweegt, en die elke vast-stelling, elk weten, in vraag stelt. Voor de wetenschap en het verstand is de vraag naar zin nutteloos: waarom denken? Waarom ons inlaten met een bezigheid die alleen onzekere en oncontroleerbare resultaten oplevert en die bovendien die resultaten altijd zelf weer vernietigt. Heeft het weten en de kennis niet juist de functie dat we ons in de wereld kunnen oriënteren, er ons thuis in kunnen voelen, er onze weg in vinden, de wereld kunnen vatten. Moeten we niet tot de orde van de dag overgaan ? Het denken is in dit opzicht altijd storend, omdat het zoals Heidegger zei precies "ausser der Ordnung" is, buiten de orde. In die zin is het inderdaad ook niet realistisch, niet gebonden aan de *'ascholia'*, aan de dringendheid van 'de orde van de dag' en de doelgerichtheid van het 'gewone' leven. Het denken ontwerkelijkt (i.p.v. te verwerkelijken), en kan wellicht in die betekenis en uitsluitend in die betekenis idealistisch genoemd worden.

In het verlengde van het onderscheid tussen weten/kennen en denken ligt het onderscheid tussen leren leren en leren denken. Leren leren verwijst naar het verwerven van een algemene bekwaamheid ingewikkelde *problemen* te begrijpen en doeltreffend in te grijpen, het is het verwerven van strategieën en methoden. Leren denken verwijst niet naar een methode, maar veeleer naar een leren luisteren, naar een 'horen' van wat de wereld (de andere) ons te zeggen heeft door en over *onze* voorstellingen of vaststellingen, ons weten en onze kennis heen.

Wanneer gepleit wordt voor een speelruimte voor het denken dan is dat geen pleidooi voor een belangloze vrijplaats, afgeschermd van het weten. Het is een vragen om aandacht voor de ruimte die *in* het weten onstaat wanneer de vraag die *wij* niet stellen opduikt: dat is de vraag naar zin, of zo men wil de normatieve vraag, de vraag naar geldigheid. Het komt er precies op aan die vraag niet toe te vertrouwen aan een 'weten', ook geen gespecialiseerd weten, omdat het geen vraag van het weten is, omdat elk weten er een ontkenning van is.

Het leren denken, vorming, is niet zozeer iets leren, maar een bepaalde wijze van leren, zoals ze bv. in openbare seminaries naar mijn mening plaats vindt. Dergelijke seminaries zijn niet vormend omdat ze in een weten en kunnen resulteren, maar omdat ze de vorm van het denken (en leren denken) zelf zijn, omdat ze een vorm zijn van het oneindige gesprek

(Blanchot) dat tegelijk onderwijs-en-onderzoek is, vragen en verantwoorden, stellen en in vraag stellen, denken en ver-denken, ordening en verstoring van de orde.

Misschien is het vandaag meer dan ooit onze plicht 'de orde van de dag' te verstoren, onze plicht wanneer we de aan-spraak van het denken 'horen'. En misschien is dat inderdaad de laatste strohalm van een idealisme dat zich weigert neer te leggen bij het dictaat van de 'werkelijkheid', dat, zoals Adorno schrijft, "Der Traum der Bildung, Freiheit vom Diktat der Mittel, der sturen und kargen Nützlichkeit" niet vervalst "zur Apologie der Welt, die nach jenem Diktat eingerichtet ist" (Theorie der Halbbildung).

De jaarlijkse bijeenkomst van het 'International Congress for School Effectiveness and Improvement': Norrköping 1993

Het ICSEI (International Congress for School Effectiveness and Improvement) is ruim vijf jaar geleden gestart als ICSE; een internationale organisatie met aandacht voor school effectiviteit. In de periode daaraan voorafgaand is de onderzoeksmatige aandacht voor factoren die scholen effectief maken snel toegenomen. Een indicatie daarvoor is dat de eerste reviews over dit thema uit die tijd stammen (bijv. Effective schools: a review door Purkey & Smith in het Elementary School Journal van 1983). Het resultaat van de toenmalige wetenschappelijke arbeid is het bekende overzicht van vijf schoolfactoren, die onderscheidend zouden zijn voor effectieve en minder effectieve scholen. Sindsdien is dit 5-factoren model aan de nodige kritiek onderhevig geweest. Zo stelt Creemers dat variantie in effectiviteit van onderwijs in veel sterkere mate bepaald wordt door instructiekenmerken dan door factoren op schoolniveau, en dat de school als zodanig er feitelijk niet zoveel toe doet. De kritiek van Scheerens richt zich op tekorten in de empirische onderbouwing van het 5-factoren model.

Al vrij snel is binnen het ICSE de aandacht verbreed tot vraagstukken op het terrein van onderwijsvernieuwing. Met de constatering dat bepaalde factoren kenmerkend zijn voor effectieve scholen is een eerste stap gezet. De volgende stap is om na te gaan of en hoe deze factoren op zinvolle wijze toegepast kunnen worden in minder effectieve scholen. Hiermee wordt de relatie tussen School Improvement en School Effectiveness aan de orde gesteld. Enkele jaren geleden is de thematiek en de naam van de organisatie dan ook uitgebreid tot ICSEI. Het samengaan van de bovenstaande twee stromingen moet vooralsnog als een 'verstandshuwelijk' getypeerd worden. Integratie wordt bemoeilijkt door verschillen in object van onderzoek, onderzoeksbenadering en visie op de school als organisatie, zoals uit analyses van bijvoorbeeld Reynolds blijkt.

Terwijl de afstemming tussen deze benadering van scholen nog heel wat werk vraagt, heeft zich op het congres in Norrköping in verschillende papers en symposia een vraagstuk aangediend dat de hiervoor geschetste verhouding tussen effectiviteit en vernieuwing (vernieuwing ten dienste van effectiviteitsverbetering) in zekere zin lijkt om te keren. Dit vraagstuk betreft een regelmatig terugkerende problematiek (althans in de westerse wereld, zie bijvoorbeeld het Swedish National Assessment Programme, in Norrköping gepresenteerd door Ekholm en collega's), en het laat zich kortweg als volgt verwoorden. Met het oog op kwaliteitsverbetering van het onderwijs en budgetbeheersing door de overheid krijgen scholen in veel westerse landen meer autonomie. Bovendien zijn er meer dan voorheen gegevens beschikbaar over de feitelijk gerealiseerde output van scholen, die als drijfveer en richtinggever moeten dienen bij het werken aan kwaliteitsverbetering. De vraag die zich voor beleidsmedewerkers, ondersteuners en beleidsgerichte onderzoekers - een omvangrijke sub-populatie van congresgangers - aandient is hoe men scholen nu zo ver krijgt dat ze daadwerkelijk iets met deze beleidsruimte en feedback van output gegevens gaan doen. De omkering in de verhouding tussen effectiviteit en vernieuwing is dat gevraagd wordt hoe gegevens over de effectiviteit van scholen een aanzet kunnen geven tot vernieuwingsprocessen in deze scholen.

Vanuit een praktisch beleidsmatig oogpunt zal het hiervoor besproken probleem de komende jaren ook in Nederland en Vlaanderen actueel worden (zie bijvoorbeeld de WRR-nota over de verzorgingsstructuur in Nederland). Het is daarom de moeite waard om het verschijnsel performance feedback naar leerkrachten en scholen in het kader van effectief onderwijs en onderwijsvernieuwing theoretisch en onderzoeksmatig nader te beschouwen. Begrippen die daartoe bruikbare handvatten bieden zijn doelmatigheidsverwachting of self-efficacy en schoolcultuur. In zijn inleiding voor het congres wees Mortimore op het grote belang van doelmatigheidsverwachting van

leerkrachten voor effectief onderwijs. In het algemeen is bekend dat performance feedback een belangrijke rol speelt bij de ontwikkeling en verandering van doelmatigheidsverwachting. Het is een alleszins redelijke aanname dat performance feedback naar leerkrachten de doelmatigheidsverwachting van deze leerkrachten in positieve of, onbedoeld, in negatieve zin zal beïnvloeden. Een vergelijkbare redenering is mogelijk voor de invloed van performance-feedback op de schoolcultuur. Een verdieping van het inzicht in de invloed van performance feedback op doelmatigheidsverwachting van leerkrachten en op de schoolcultuur, en daarmee op het leren van leerkrachten en van de school als organisatie, kan een zinvolle bijdrage leveren aan het vergroten van de praktische bruikbaarheid van allerlei assessment-gegevens (periodieke peilingen, leerlingvolgsystemen, etc.) die de komende tijd in steeds grotere getale beschikbaar zullen komen.

De rol van performance feedback op kwaliteitsaspecten van het onderwijs hoeft niet als een geïsoleerd probleem beschouwd te worden. Aansluiting is mogelijk bij diverse organisatietheoretische en organisatiekundige inzichten, bijvoorbeeld inzichten in de rol van cognities in het functioneren van organisaties (organisaties als informatieverwerkende systemen), of inzichten in het leren van organisaties en kwaliteitszorg op de werkvloer. Een theoretische implicatie van aandacht voor performance feedback voor het gangbare schooleffectiviteit-model is dat de eenzijdigheid wordt doorbroken in de relaties tussen schoolkenmerken en output-indicatoren. In een theoretisch model waarin de bedoelde feedback-processen zijn opgenomen gaat de aandacht niet alleen uit naar de invloed van schoolkenmerken op schooleffectiviteit, maar ook naar de invloed van informatie over de schooleffectiviteit op van belang geachte school- en leerkrachtkenmerken.

In het 'invited paper' van Fink en Stoll, een paper waarin wordt gereflecteerd op de ervaringen die in de Halton Board of Education (omgeving Toronto) zijn opgedaan met het implementeren van effectieve-schoolkenmerken in scholen, wordt het hiervoor geschetste probleem als volgt gesteld: "How to balance the system's emphasis on teacher empowerment with an approach to accountability which is seen as professionally enabling?". Volgens hen was een werkwijze met goede resultaten om scholen uit te nodigen om van beschikbare assessment-instrumenten gebruikt te maken. Gegevens over de school behoorden alleen aan de school toe. Op het niveau van het district werd aan het bestuur en de pers gerapporteerd. Door deze benadering is de behoefte aan ondersteuning bij assessment binnen scholen sterk toegenomen. Een meer algemene conclusie van Fink en Stoll is dat het de moeite waard is om schooleffectiviteit en school improvement met elkaar te verbinden, maar dat het hierbij om een langzaam proces gaat dat om veel geduld en volharding vraagt, omdat het veranderen van scholen een verandering inhoudt van de schoolcultuur.

J. Imants, Rijksuniversiteit Leiden

Boekbespreking

C. Vreugdenhil

De Führungslehre van Peter Petersen. Historisch-pedagogisch tekst- en context-onderzoek van de onderwijspedagogiek van het Jenaplan

Academisch proefschrift Universiteit van Amsterdam.
Wolters-Noordhoff bv, Groningen, 1992
374 + 158 p. + bijlagen ƒ 85,– ISBN 90 01 91982 0

Op 18 juni 1992 promoveerde Vreugdenhil aan de Universiteit van Amsterdam op een studie naar de grondslag en de inhoud van een van Petersens hoofdwerken, de *Führungslehre* (promotor: L. Dasberg). Het is, ruim tien jaar nadat in Nederland de eerste monografie over de Petersen-pedagogiek verscheen van de hand van 'buitenstaanders' (Imelman, Jeunhomme & Meijer, 1981), goed dat zich nu ook een 'insider' over Petersens thematiek heeft gebogen.

De studie omvat twee delen: een verslag van het onderzoek en een band met aantekeningen, een bibliografie, personen- en zakenregisters en bijlagen (o.m. bestaande uit recensies van Petersens 'Lehre' en, interessant, een vergelijking tussen de tekst van de eerste en de, deels van nazi-vriendelijke taal gekuiste, tweede druk van resp. 1937 en 1949).

In 1 (Inleiding) duidt de auteur de centrale betekenis van de 'Führungslehre' binnen het totale werk van Petersen aan. Hij heeft drie redenen om zijn onderzoek te verrichten: *a* voor 'Jenaplanmensen' is het boek nogal ontoegankelijk, wat jammer is want *b* kennis van het boek zou nu juist in discussies rondom het Jenaplan van pas kunnen komen; en bovendien ligt er *c* mogelijk ook nog een actuele waarde in het onderwijsconcept van het boek. Daarnaast besteedt hij aandacht aan in de jaren tachtig in Nederland gevoerde discussies rondom het Petersen-concept, en aan het sinds 1984 ook in Duitsland gevoerde debat.

De standpunten van critici van het Petersen-concept worden ook kritisch beoordeeld door Vreugdenhil zelf. Hij vindt dat er 'goede redenen' zijn om een nieuw onderzoek te doen. Door de 'abstractiegraad van begrippen te verminderen' wil hij proberen tot een 'gedifferentieerde invulling van begrippen' te komen (p. 37).

Ten slotte omschrijft de auteur zijn werkwijze: een communicatietheoretische aanpak waarbij in het tekstonderzoek de (oorspronkelijke) auteur, tekst, lezer en context op elkaar zullen worden betrokken (p. 40 e.v.).

De opbouw van de studie is als volgt.

Omdat de 'Führungslehre' het hoofdthema van het onderzoek is, staat vooral de periode waarin Petersen de desbetreffende conceptie denkend en handelend ontwikkelde centraal. Grofweg gaat het om 1920-1938. (Hij begon vermoedelijk in het begin van de jaren dertig met schrijven; p. 360). Van die jaren wordt de maatschappelijke context geschetst (hoofdstuk 2) en worden het leven en werk in onderling verband besproken (hoofdstuk 3). De meest opvallende onderwerpen in deze twee hoofdstukken vormen 'Weimar' (het culturele leven, de wijze waarop de democratie zichzelf uitprobeerde, de vanaf 1933 in Nazi-kringen ontwikkelde opvoedings- en onderwijsideologie, de 'omstreden opbouw van een eigen instituut' voor 1933 en daarna (in Jena). Uiteraard wordt aan de perioden voor en na het in extenso behandelde tijdperk ook aandacht besteed.

Na de geparafraseerde samenvatting van de 'Lehre' in hoofdstuk 4 volgen in de hoofdstukken 5 tot en met 7 reacties van tijdgenoten op het werk, een beschrijving van twee opvoedingsconcepten van Sturm en van het pedagogisch relevante socialistische gedachtengoed van Reichwein (in 1944 geëxecuteerd na een poging contact te leggen met communistische verzetsgroepen). In hoofdstuk 8 volgt dan een vergelijking tussen de concepties van Petersen, Reichwein en Sturm. Het eerste concept 'kenmerkt zich ... door het primaat van de pedagogische doelen, het uitgaan van absolute waarden, werken met thema's uit de leefwereld van de kinderen, het verwerken van modern-psychologische inzichten (van o.a. de typoloog Jaensch, op wiens uitgesproken racisme Vreugdenhil overigens vergeet te wijzen; JDI) ..., de plicht tot toewijding aan het volk ...' enz. Het onderscheidt zich in deze punten niet van een der concepties van Sturm of van dat van Reichwein, aldus Vreugdenhil. Wel verschilt het van deze dáárin dat kinderen een belangrijke invloed toegestaan wordt 'in de vormgeving van ... relaties en in de keuze van leerinhouden, en door nauwelijks te hechten aan tucht ...' (p. 361).

Verrassenderwijs vergelijkt Vreugdenhil Petersens ideeën ook met die van Van Parreren

en Dasberg. Hij meent overeenkomsten met de opvatting van Van Parreren te zien: beiden 'werken met de grondprincipes van een vak', zijn gericht op 'leerprocessen en minder op -resultaten', kennen 'de waarde van initiatief van kinderen', zien 'de rol van de leraar als begeleider'. De overeenkomsten met Dasberg hebben betrekking op 'liefde voor het kind, het nut van een 'waarden'volle opvoeding, integratie van opvoeding en onderwijs, brede persoonlijkheidsontwikkeling, vakkenintegratie en de school die open staat voor ontwikkelingen in de maatschappij ...'. Dasberg en Van Parreren verschillen echter met Petersen van mening, aldus de auteur, op het punt van zijn 'grondslag': de beide hedendaagse Nederlanders kiezen voor 'een rationele, cultuurhistorische analyse om kennis over mens en werkelijkheid te genereren' (de grondslag van de 'Führungslehre' is eerder irrationeel en in ieder geval, aldus de onderzoeker 'eenzijdig' - want 'de historische dimensie van cultuur' wordt verwaarloosd) (p. 361 e.v.).

Vreugdenhil meent ten slotte dat de 'Lehre' 'niet als een nationaal-socialistisch onderwijsconcept' moet worden aangemerkt. Integendeel, ze kan 'vermoedelijk ... als een rivaliserende theorie met die van een aantal nazi-pedagogen worden beschouwd', en wel op het terrein van het mensbeeld en de maatschappij-opvatting. Ondertussen zal de 'Führungslehre' 'voor huidige en toekomstige lezers moeilijk toegankelijk blijven. Dat wordt vooral veroorzaakt door de metafysische grondslag ervan'. Desalniettemin kunnen het Petersen-concept (zowel als Sturms eerste, niet nationaal-socialistische, en Reichweins concept) bijdragen leveren 'aan de huidige discussie over pedagogisering van het onderwijs', omdat ze verschillende uitwerkingen laten zien 'zonder daarbij in moraliseren te vervallen'.

In hoofdstuk 9 wordt nog eens aandacht besteed aan de gevolgde onderzoeksmethode, die volgens de schrijver berust op inzichten uit de geschied- en de literatuurwetenschap.

Het mag een prestatie genoemd worden dat iemand naast zijn toch al veel tijd en aandacht opeisende 'gewone' werk (Vreugdenhil werkt bij het Algemeen Pedagogisch Studiecentrum) zo'n indrukwekkende en grondige studie heeft verricht. Niet alleen de tekst (die reeds rijk aan data is) maar vooral ook het aan data overrijke tweede deel getuigen van grondigheid en geduld - vandaag de dag schaars wordende deugden.

Het is dan ook, juist tegen de achtergrond van deze waardevolle kanten van de studie, jammer dat de auteur zich heeft laten verleiden tot het schrijven van een weinig diepgravende, meer descriptieve dan argumentatieve tekst die van dubbelzinnigheden, tegenspraken, en soms halve waarheden niet is vrij te pleiten. Of is er in het laatste geval geen sprake van opzet, doch gaat het om gebrek aan cultuurhistorische eruditie? Zo is Vreugdenhil bijvoorbeeld nogal eenzijdig in het volgende fragment.

Op p. 61 schetst hij hoe een stortvloed van nationaal-socialistische wetten na maart 1933 het nieuwe rijk overspoelt. Tegelijkertijd 'lopen vele bekende persoonlijkheden over naar de nationaal-socialisten'. Waarom? Omdat zij, zo citeert hij Thamer (1986), de nazistische mentaliteit gingen waarderen; de mentaliteit van de onderwerping 'aan de geborgenheid in de gemeenschap', van 'de roes van massabijeenkomsten', de mentaliteit van 'de bereidheid in een utopie te geloven en offers te brengen' en van 'het verlangen naar eenvoud en onvoorwaardelijkheid'. Wie noemt Vreugdenhil dan? Van de kunstenaars noemt hij er maar twee: Gründgen (terecht!) en ... Furtwängler, een uiterst discutabel voorbeeld, zoals iedere muziekhistoricus weet die de tegendraadsheid van deze dirigent kent. 'Gerenommmeerde tegenstanders uit de voorgaande jaren', zo gaat hij voort, 'beginnen te twijfelen: *zou de kern van de beweging dan toch juist zijn* (curs. JDI)? Zo is Adorno onder de indruk van de gedichten van Baldur von Schirach... en Thomas Mann betreurt het niet dat de parlementaire democratie verdwenen is en dat joden uit het rechtsapparaat worden verwijderd' (p. 61). Wat rechtvaardigt het noemen van Adorno, de man die in 1934 al vluchtte en wiens anti-nazisme tot diepgaand onderzoek naar de fascistoïde persoonlijheid heeft geleid? En waarom van Mann dit verteld, als hij al geëmigreerd *is* en overigens voor 1933 al via publikaties heftig had gewaarschuwd tegen het nazisme? Het noemen van deze namen uitsluitend en alleen in dit verband is tendentieus, zeker als Vreugdenhil hun de boven gecursiveerde vraag in de mond geeft. (Vreugdenhils opmerking dat Manns werk 'zeker tot 1939' 'is toegelaten', wordt in de Winkler Prins (dl. 12, 1971) weersproken.)

Tot de grootste tekortkomingen is in dit verband te rekenen dat Vreugdenhil zijn contextonderzoek voornamelijk beperkt tot het bij elkaar brengen van feiten die moeten tonen dat het allemaal erg begrijpelijk is dat Petersen schreef wat hij schreef. Het 'andere Duitsland' (van bijvoorbeeld de steeds meer tegen de tijdgeest opponerende collega-pedagoog Litt, de pedagogen 'in Exil', sociale wetenschappers, schrijvers en kunstenaars als Kurt Lewin, Brecht, Tucholsky, Kästner, de Frankfurters en vele anderen) wordt niet in het geding gebracht.

Helemaal bont maakt Vreugdenhil het als hij (zie boven) Petersens conceptie (inclusief de metafysica dus) vergelijkt met de betrekkelijk onsystematische opvattingen van Dasberg en met Van Parrerens leerpsychologie. Zonder enige aarzeling gaat hij voorbij aan het gegeven dat het in deze drie gevallen om volstrekt onvergelijkbare theoretische invalshoeken gaat (p. 297 e.v.; p. 361 e.v.).

Maar niet alleen met namen en met 'taalspelig' van elkaar te onderscheiden theoretische benaderingen wordt onverantwoord gegoocheld, ook met de interpretatie van de door hem wel terecht in ogenschouw genomen contextuele feiten. De 'Führungslehre' vergelijkend met concepties van Sturm en Reichwein concludeert Vreugdenhil o.m. dat Petersens schoolconceptie een redelijke mate van 'onafhankelijkheid van de heersende machten' vraagt. Waaruit leidt hij dat af? Dááruit dat de 'opvoedingswetenschappelijke grondslag van de Führungslehre stamt uit de tijd van 'Weimar" (p. 296). De lezer moet nu kennelijk denken: o, 'Weimar', dat is democratisch, en: 'Weimar' is de context van Petersens schoolconceptie, dus: vandaar die 'onafhankelijkheid van de heersende machten'. Maar Vreugdenhil vergeet te zeggen dat Petersens opvoedingswetenschappelijke grondslag, zijn *Allgemeine Erziehungswissenschaft* (1924), juist eerder uitdrukking van een *anti*-Weimar-geest is. Daarin wordt voor het eerst systematisch aan sociaal-psychologische verschijnselen als 'massa' en 'gemeenschap' aandacht besteed, en wordt 'democratie' ernstig gelaakt. (Het inconsistente is dat Vreugdenhil elders weer wel aangeeft dat Petersens conceptie anti-democratisch is.)

Petersens pogingen om na 1933 zijn teksten beter te laten aansluiten bij nazistisch gekleurde taal laat de teneur van zijn metafysica en pedagogiek onverlet. Als Vreugdenhil ten allerleste concludeert dat de 'Lehre' 'qua grondslag en uitwerking niet als een nationaal-socialistisch onderwijsconcept' kan worden aangemerkt, en dat ze, in tegendeel, eerder gezien moet worden 'als een (daarmee) rivaliserende theorie' (p. 362), is dat echter maar zeer ten dele waar. 'Waar', omdat het inderdaad niet om een expliciet nazistische pedagogiek gaat; maar niet waar in zoverre die pedagogiek in haar centrale begrippen en in haar gehele strekking verwant is aan een aantal typisch Duits-idealistische denkvormen waartoe ook het nazistisch discours is te rekenen. Het gekke is dat Vreugdenhil dit, in een adem door, in concreto ook erkent. Hij zegt namelijk dat Petersens conceptie 'rivaliseert', omdat er sprake is van een bepaalde gemeenschappelijkheid in doelstelling en taakveld, o.a. tot uiting komend in het 'völkische', het 'nordische', 'het organische denken en het irrationele.' (Deze conclusie is ook in overeenstemming met onze eigen, destijds echter door vele Jenaplanners onacceptabel gevonden, conclusie: Imelman e.a., 1981.) 'Op fundamentele aspecten is ze echter', aldus Vreugdenhil, 'aan die van de Nazi's tegengesteld, zowel in mensbeeld (altruïsme), als in maatschappijopvatting (van onderop, met zelfbestuur). *Dat maakt haar rivaliserende karakter uit'* (p. 362/3; curs. JDI).

Nu spreekt Petersen, naar ik meen, nergens van 'altruïstisch mensbeeld'. (In Vreugdenhils onderzoek blijft, ook na raadpleging van het register, de herkomst van 'altruïsme', 'altruïstisch' duister.) Wel rept Petersen wel eens van 'humaniteit', maar dan in een zeer abstracte zin van 'alle mensen delen in menselijkheid'. De echte opvoeding speelt zich echter niet in de sfeer van 'algemene menselijkheid' maar in die van het gezin, de familie, de gemeenschap af. Gemeenschappen, zo spreekt Vreugdenhil Petersen na, vormen een noodzakelijke voorwaarde voor elke individuele ontwikkeling. Deze afhankelijkheid van de gemeenschap roept de verplichting van dienst aan de gemeenschap op (zegt Vreugdenhil, al even juist). En in het verlengde daarvan spreekt hij ineens over een 'altruïstische theorie' (p. 266). We moeten deze evocatieve misleiding natuurlijk niet serieus nemen. Petersens gemeenschapsfilosofie is en blijft in overeenstemming met wat ook nazistisch taalgebruik onder 'gemeenschap' verstaat

(inclusief de morele connotaties van dienst en anti-democratische saamhorigheid ervan).

Iets dergelijks valt er te zeggen over het tweede door Vreugdenhil 'rivaliserend' genoemde aspect van Petersens conceptie, de maatschappij-opvatting (die zelfbestuur van onder-op inhoudt). Weliswaar groeide het Derde Rijk snel uit tot een dictatuur, maar het nazistisch gekleurde jargon kent zeer duidelijk de geromantiseerde aandacht voor de Heimat, de grond, bloed en bodem, kortom: het gebied van de zichzelf besturende gemeenschap waarin zich 'natuurlijke' leiders voordoen. De 'van-onderop-ideologie' behoort tot corporatistisch gedachtengoed en past als zodanig in de nazistische denkwereld (al trokken de politieke delinquente kopstukken destijds al snel andere politieke consequenties, namelijk die van de inrichting van een dictatuur). Nergens is mijns inziens het natuurlijk leiderschap zo evident herkenbaar als in het Petersens theorie typerende idee 'leiderschap'. (Volgens Petersen verstaat de leider de geest van de gemeenschap en de tijd en helpt hij zo het telos van de geschiedenis van de gemeenschap te realiseren).

Het zogenaamde altruïsme- en het zelfbestuur-aspect van Petersens conceptie, die volgens Vreugdenhil deze pedagogiek zouden doen rivaliseren met nationaal-socialistische concepties, zijn dus direct al te ontmaskeren. Die aspecten (de term 'altruïsme' kan dat niet verhullen) behoren namelijk eveneens tot de kenmerken van het Duits-idealistische denken, waarvan ook nazistische programma's etc. zich bedienen.

Vreugdenhil acht de 'Führungslehre' 'voor huidig gebruik alleen geschikt als de grondslag ervan opnieuw wordt doordacht'. In 1981 werd dat door anderen duidelijker gezegd: men kan de metafysica van Petersen maar beter vergeten en het praktische schoolorganisatie-model op zichzelf nemen. Jenaplanners kunnen daarin voldoende 'grondslag' vinden om nog vele jaren voort te gaan met het stichten en verbeteren van hun, in vele opzichten waardevolle, onderwijspraktijk.

Vreugdenhils studie wordt, bij nader inzien, ontsierd door een verhullende, sterk op klank- (en niet op begrips)associatisch berustende taal. Tegenspraken weet hij niet te vermijden. Ook de vele stellingen waarmee een en ander gelardeerd is, doen afbreuk aan de kwaliteit van het onderzoek. Vrijwel elke stelling laat zich op het punt van meerzinnigheid en evocatief taalgebruik, van het bij elkaar vegen van kwalitatief verschillende begrippen en soms een tegenspraak, bekritiseren. (Een volstrekt willekeurig gekozen voorbeeld: stelling 13 op p. 268 luidt: 'Het karakter van de opvoedingswetenschappelijke theorie van Petersen is organistisch-realistisch, actionistisch, sociaal-personalistisch, altruïstisch, optimistisch, historistisch, irrationeel met rationeel-technische consequenties, en pedagogisch-ethisch'. Niet alleen komen hier in een zin onderling onverenigbare begrippen voor, ook met de rest van de studie bestaan er tegenspraken. Is Petersen hier getekend als een historisch ad absurdum denkende 'historist', elders wordt hem, o.a. in vergelijking met Dasberg, een onhistorische instelling verweten; zie eerder.)

Een enkel woord nog over de gevolgde onderzoeksmethodiek. Vreugdenhil beroept zich, bij de verantwoording van de door hem gevolgde methode, voornamelijk op *inleidende* literatuurwetenschappelijke publikaties inzake receptie. Van het toepassen van communicatietheoretische methodologische inzichten is, anders dan hij zegt, geen sprake. De traditie waarin Petersens denken staat, wordt - los van het aanduiden van enkele referenten die Petersen zelf aangeeft - niet in beschouwing genomen. De context-analyse schiet, naast de boven genoemde kritiek, ook op dit punt te kort.

De opbouw van de studie - toch ook methodisch relevant - is op onderdelen niet plausibel. Zo kennen de daartoe geëigende hoofdstukken geen entree in de sinds ongeveer tien jaar gevoerde debatten over de behandelde onderwerpen; de lezer moet het doen met de nogal summiere descriptie van discussies in het inleidende hoofdstuk. Elke nadere beschouwing van de uitgewisselde argumenten ontbreekt aldaar. En het, op zich natuurlijk schitterende, materiaal van de kritieken van tijdgenoten wordt ook alleen maar weergegeven, soms gecategoriseerd, maar nooit in een argumementatief betoog nader aan de orde gesteld.

Al met al gaat het om een in wetenschappelijk opzicht teleurstellend boek. De waarde ervan ligt in het verzamelde materiaal en de wijze waarop dit is opgenomen in de tekst en is bijgevoegd. Dat maakt deze dissertatie per slot van rekening wel tot een 'must' voor iedereen die in

dit stuk van de geschiedenis van het Jenaplan is geïnteresseerd.

J. D. Imelman

Literatuur

Imelman, J. D, J. M. P. Jeunhomme en W. A. J. Meijer (1981). *Jenaplan. Wel en wee van een schoolpedagogiek*. Nijkerk: Intro

Petersen, P. (1924), *Allgemeine Erziehungswissenschaft*. Leipzig

Thamer, H.-U. (1986). *Verführung und Gewalt. Deutschland 1933-1945*. Band 5 van *Die Deutschen und ihre Nation*. Berlin: Siedler Verlag

R. H. Hofman

Effectief schoolbestuur. Een studie naar de bijdrage van schoolbesturen aan de effectiviteit van basisscholen

Academisch proefschrift, RION-monografieën onderwijsonderzoek nr. 13, Groningen, 1993.
299 pagina's + bijlagen. *f* 65,– ISBN 90-6690-344-9.

In haar proefschrift *Effectief schoolbestuur* onderzoekt Hofman of schoolbesturen via hun bestuursbeleid een bijdrage leveren aan leerprestaties van leerlingen. Haar proefschrift is één van de eerste Nederlandse onderzoeken waarin wordt geprobeerd een relatie te leggen tussen het beleid van schoolbesturen en leerlingprestaties.

Het onderzoek heeft betrekking op openbare en bijzondere schoolbesturen van het basisonderwijs. Hofman gaat uit van twee probleemstellingen. In de eerste probleemstelling stelt zij zich de vraag welke organisatiekenmerken van schoolbesturen optreden als determinanten van bestuurlijke activiteit. In de tweede probleemstelling is de vraag welke bijdrage schoolbesturen via het bestuursbeleid leveren aan de effectiviteit van het onderwijs in de door hen bestuurde school. Het bestuursbeleid is in de eerste probleemstelling dus de afhankelijke variabele, in de tweede juist de onafhankelijke variabele.

In de eerste probleemstelling zijn vier clusters van bestuursfactoren opgenomen. De 'context-factor' is de denominatie van het bestuur. De 'contingentiefactoren' zijn bestuurlijke schaalgrootte, mate en soorten van externe bestuurlijke ondersteuning en de invloed van interne en externe groeperingen op bestuursbeslissingen. Bij de 'structuurkenmerken' gaat het onder meer om interne bestuursdifferentiatie, de verdeling van beslissingsbevoegdheid en deskundigheid van schoolbesturen. Het 'outputcluster bestuursbeleid' ten slotte, heeft betrekking op de activiteiten van schoolbesturen op zes handelingsdomeinen (interne en externe contacten, personeels-, onderwijs-, organisatie- en materieel en financieel beleid) en de taakopvatting van schoolbesturen.

Bij de tweede probleemstelling stelt Hofman dat schoolbesturen leerlingprestaties indirect kunnen beïnvloeden, namelijk via het realiseren van effectiviteitsbevorderende schoolkenmerken. Overigens relativeert zij wel de omvang van de bijdrage die besturen leveren aan de effectiviteit van scholen. Die bijdrage kan hoogstens een gedeelte zijn van de 12% tussenschoolse variantie, die in Nederland als gemiddeld percentage wordt geschat. Ook bij deze probleemstelling spelen vier clusters van factoren een rol. Het bestuur is in het 'contextcluster' geplaatst. Hierbij gaat het om het bestuursbeleid, contingentiefactoren, structuurkenmerken en vier typen besturen (uitgewerkt op basis van de resultaten van de eerste probleemstelling). De 'inputvariabelen' zijn de mate waarin sprake is van extra formatie (in hoeverre heeft de school veel leerlingen uit laag sociaal-economische milieus en etnische minderheidsgroepen), percentage tijd dat de schoolleider besteedt aan de ontwikkeling van het schoolbeleid, frequentie van teamoverleg en de mate waarin leerkrachten participeren in besluitvorming. De 'procesvariabelen op schoolniveau' zijn gebaseerd op het vijffactoren-model van schooleffectiviteit. Dit zijn onderwijskundig leiderschap, prestatiegerichtheid èn positieve verwachtingen, evaluatie van het onderwijs op schoolniveau, nadruk op basisvaardigheden, mate van orde en regelgeving en evaluatiegericht leerlingbeleid. Als 'outputvariabelen' zijn taal- en rekenprestaties (Cito-toets) en schoolbeleving (schoolbelevingsschaal) van leerlingen in groep 8 gebruikt, gecontroleerd voor de aanvangsprestaties in groep 7 en voor individuele verschillen in instroom (IQ, sekse, etniciteit en milieu).

Het onderzoek bestond uit een case-study en een survey. In de vergelijkende gevalstudie heeft zij 14 schoolbesturen onderzocht met een bijbehorende school. De resultaten gebruikt zij met name om de hypothesen van het survey nader te specificeren. In het survey zijn deze hypothesen

getoetst. In totaal hebben 146 besturen in combinatie met één onder dat bestuur ressorterende basisschool meegedaan aan het survey. Van de besturen is 32% openbaar, 39% rooms-katholiek, 25% protestants-christelijk en 4% bijzonder-neutraal. In een groot deel van de analyses laat Hofman de bijzonder-neutrale besturen buiten beschouwing, omdat dit maar een klein aantal betreft.

De resultaten van het survey met betrekking tot de eerste probleemstelling laten zien dat besturen het meest actief zijn op het terrein van communicatie, vooral communicatie met interne geledingen. Verder blijkt onder meer dat bijzondere besturen kleinschaliger en minder complex van aard zijn dan openbare besturen en hun bestuursbeslissingen meer worden beïnvloed door de eigen schoolgeledingen.

Een analyse van de samenhang tussen contingentiefactoren en structuurparameters, (waarbij zij tot een indeling in vier bestuurstypen komt) laat zien dat de factor bestuurlijke schaalgrootte in feite meer gewicht in de schaal legt dan denominatie. Er zijn niet alleen verschillen tússen openbare en bijzondere besturen, maar ook binnen de verschillende denominaties. Verder blijkt dat de vier bestuurstypen significant verschillen op het terrein van onderwijsbeleid en financieel-materieel beleid. Maar als gekeken wordt naar de invloed van iedere variabele afzonderlijk, dan zien we een ander resultaat. Zo blijken schoolgrootte en urbanisatiegraad een rol te spelen. Besturen van grote scholen ontplooien meer onderwijsgerichte bestuursactiviteit dan besturen van kleine scholen. Besturen in stedelijke gebieden ontplooien meer personeelsbeleid en juist minder financieel-materieel beleid. Ook denominatie is van invloed: openbare besturen zijn minder actief op het financieel-materieel beleid dan bijzondere besturen. Verder is er sprake van meer bestuursbeleid als, onder meer, de invloed van interne groeperingen op het bestuursbeleid relatief groot is, de invloed en ondersteuning van externe instanties relatief groot zijn en het bestuur meerdere schooltypen bestuurt.

Afzonderlijke regressie-analyses voor openbare en bijzondere besturen laten zien dat er sprake is van een contrast tussen openbare besturen en rooms-katholieke besturen. Openbare besturen scoren op alle vier de bestuursdomeinen relatief laag en rooms-katholieke besturen juist op alle vier relatief hoog. De protestants-christelijke besturen voeren relatief minder taken uit op onderwijs- en personeelsbeleid. Alleen op financieel-materieel terrein is er sprake van een significant verschil tussen openbare en bijzondere besturen in zijn geheel.

Voor de analyses van de tweede probleemstelling construeert Hofman effectiviteitsindices per school in termen van reken- en taalprestaties en schoolwelbevinden, onder controle van instroomkenmerken. Deze indices vormen de criteriumvariabelen in het onderzoek en bieden de mogelijkheid de scholen te rangschikken van laag tot hoog effectief. Per type bestuursbeleid stelt Hofman vervolgens de samenhang vast met de mate van schooleffectiviteit op de drie outputmaten.

Het is niet zo dat scholen waarbij de besturen een intensief onderwijsbeleid ontplooien, effectiever zijn. Maar de scholen waarbij de besturen een relatief intensief financieel-materieel beleid ontplooien, zijn wel effectiever dan de scholen waarbij de besturen dit niet doen. Dit geldt overigens alleen voor reken- en taalprestaties. Verder blijkt dat denominatie consistent optreedt als voorspeller van reken- en taalprestaties: openbare scholen presteren minder goed dan katholieke scholen.

In de laatste stap staat de vraag centraal naar de determinanten van schooleffectiviteit: is er nog een bijdrage van schoolbesturen nadat voor invloed van (met name) schoolvariabelen is gecontroleerd? In het algemeen worden verschillen in cognitief en niet-cognitief functioneren voor het grootste deel verklaard door denominatie en stimuleringsschool (opgevat als de aanwezigheid van extra formatieplaatsen aan de school). Zowel met betrekking tot reken- als taalprestaties blijken betere resultaten behaald te worden in rooms-katholieke scholen dan in openbare scholen. Bovendien wordt voor beide outputmaten een contexteffect gevonden: in basisscholen waar de interne geledingen relatief veel invloed uitoefenen op bestuursbeslissingen, liggen de reken- en taalprestaties hoger. Met betrekking tot de schoolbeleving blijkt dat deze in protestants-christelijke scholen minder positief is dan in rooms-katholieke scholen. Ook hier treedt een contexteffect op: in scholen waarbij de besturen relatief veel personeelsbeleid ontplooien, hebben leerlingen een negatievere schoolbeleving dan in scholen waar besturen

weinig personeelsactiviteit ontplooien.

De verwachting dat de bijdrage van besturen vooral toe te schrijven zou zijn aan het bestuursbeleid, wordt niet bevestigd. Er treedt een ander contexteffect op, namelijk de invloed van interne geledingen. Dit effect neemt Hofman apart onder de loep en zij gaat na welke interne geledingen invloed uitoefenen op de bestuursbeslissingen en relateert dit aan de drie effectiviteitsmaten. Er blijkt een sterke samenhang te bestaan tussen de invloed van schoolleider en groepsleerkrachten op bestuursbeslissingen en de cognitieve outputmaten. Ook constateert Hofman een positieve samenhang tussen de invloed van ouders en ouderraad en cognitieve en affectieve outputmaten. Voor de medezeggenschapsraad gaat dit niet op. Tevens hangt de invloed van niet-onderwijzend personeel positief samen met de effectiviteit van basisscholen op het cognitieve domein.

Ze kijkt ook naar de invloed van interne geledingen op bestuursbeslissingen in openbaar en bijzonder onderwijs afzonderlijk. Er blijkt op dit punt vooral een contrast tussen openbare en katholieke besturen: katholieke besturen kennen hun interne geledingen meer invloed toe dan openbare besturen, maar ook dan protestants-christelijke besturen.

Het proefschrift van Hofman betreft een interessant en complex onderzoek. Dat wil niet zeggen dat ik het altijd met evenveel plezier gelezen heb. Dat komt vooral door de rommelige opzet en de inconsistente opbouw van het boek.

Verder is een aantal bezwaren aan te voeren tegen de operationalisatie en meting van variabelen. Een voorbeeld hiervan is het begrip 'cliëntgerichtheid', dat een centrale rol speelt in de conclusies van het onderzoek. Hofman stelt: "Eén factor blijkt uiteindelijk relevant te zijn, namelijk de invloed van geledingen in en rond de school op het bestuursbeleid. In feite zou hier gesproken kunnen worden van de 'cliëntgerichtheid' van het schoolbestuur" (p. 289). Bij deze hoofdconclusie wil ik twee kanttekeningen plaatsen. In de eerste plaats heb ik kritiek op de operationalisatie en betekenis van het begrip 'cliëntgerichtheid'. Hofman omschrijft dit als beleidsinvloed vanuit de school (p. 192). Als we bedenken dat in dit onderzoek de 'invloed van interne geledingen' niet alleen betrekking heeft op ouders, maar ook op schoolleider en team, dan is de term cliëntgerichtheid heel opmerkelijk. Dat zou namelijk betekenen dat het personeel aan de school tevens cliënt van de school is. Maar het begrip impliceert dat het gaat om de consumenten van het onderwijs: de ouders en de leerlingen. Dat in Hofmans studie ook het personeel van de school onder 'cliëntgerichtheid' valt, vind ik dan ook verwarrend.

In de tweede plaats stelt Hofman de mate van invloed van interne geledingen vast via de mate van invloed zoals het bestuur dit percipieert en niet zoals de interne geledingen dit zelf percipiëren. Of de interne geledingen zichzelf net zoveel invloed toekennen als de besturen dit doen, is natuurlijk maar de vraag.

Een ander voorbeeld is dat het bestuursbeleid niet alleen gemeten wordt door middel van de activiteiten van schoolbesturen, maar ook door middel van hun taakopvatting. Dit is een opvallende keuze. In de eerste probleemstelling is het bestuursbeleid immers de afhankelijke variabele, terwijl taakopvatting een onafhankelijke variabele is: "Het zou met name deze taakopvatting van het schoolbestuur zijn die de bestuursactiviteit op een aantal verschillende taakgebieden zou sturen" (p. 47). Het is mijns inziens dan ook onterecht om taakopvatting als operationalisatie van de afhankelijke variabele bestuursbeleid te hanteren.

Een ander kritiekpunt betreft de verantwoording van de methode. Het is merkwaardig dat Hofman bij het survey geen expliciete aandacht besteedt aan de methode van dataverzameling, betrouwbaarheid en validering, terwijl zij hier bij de case-study een aparte paragraaf aan wijdt. Ze beschrijft wel hoe de instrumenten voor case-study en survey ontwikkeld zijn, maar vermeldt niet specifiek welke data-verzamelingstechniek gebruikt is in het survey. Zo is niet duidelijk wie nu precies de vragenlijsten hebben ingevuld. In de Engelstalige samenvatting valt alleen te lezen dat vragenlijsten naar besturen en schoolleiders zijn gestuurd.

Een grote verdienste van het onderzoek ligt in de gewaagde opzet ervan: voor het eerst is de causale keten van het beleid van schoolbesturen naar de uiteindelijke leerlingprestaties onderzocht. Dit is uniek in Nederland. Tevens is getracht om twee verschillende theoretische perspectieven in één onderzoek te koppelen: een contingentie-theoretische benadering van bestuur en beheer van scholen en het onderzoek

naar schooleffectiviteit. Bovendien heeft Hofmans onderzoek een waardevol inzicht opgeleverd: schoolbesturen doen er wel degelijk toe. "Indien we daarbij de marges van schoolbesturen in ogenschouw nemen, namelijk het gegeven dat de meeste variantie op het niveau van de school ligt, dan is bovenstaand resultaat des te opmerkelijker", aldus Hofman (p. 289).

Tegelijkertijd ligt hier de beperking van het onderzoek. Eén van de belangrijkste uitkomsten van het onderzoek is immers, dat er een samenhang bestaat tussen de mate van cliëntgerichtheid van het bestuur en de effectiviteit van een school. Bovendien blijkt dat rooms-katholieke besturen niet alleen cliëntgerichter zijn dan openbare besturen, maar ook dan protestants-christelijke besturen. Met andere woorden: er is niet alleen een verschil tússen openbare en bijzondere, maar ook binnen bijzondere besturen. Wat nu precies de rooms-katholieke besturen zo bijzonder maakt ten opzichte van de protestants-christelijke besturen, is een vraag die nog steeds onbeantwoord is gebleven. Deze kwestie behoeft, zeker gezien de huidige discussie over de bestuursvorm, nader onderzoek.

A. M. den Haan

J. M. Sijtstra

Doel en inhoud van het taalonderwijs. De ontwikkeling van een model voor domeinbeschrijvingen van taalonderwijs

Academisch proefschrift Katholieke Universiteit Brabant. Cito, Arnhem 1991

IX + 201 pagina's + bijlagen

Inleiding

Sijtstra behandelt een belangrijk onderwijskundig probleemveld: de doelstellingenproblematiek. Haar onderzoeksvraag geldt niet wat leerlingen zouden moeten leren of wat ze leren, maar de structuur van het doelstellingendomein. Het gebied dat zij in kaart probeert te brengen is ruim: het taalonderwijs. Daaronder rekent zij het onderwijs in het Nederlands (moedertaalonderwijs), het onderwijs in het Nederlands als Tweede Taal en het moderne vreemde talenonderwijs, zowel in het basis- als het voortgezet onderwijs.

Directe aanleiding voor de studie van Sijtstra (hierna JS) is een CITO-project 'itembank', dat een grote hoeveelheid toetsitems voor het taalonderwijs zou moeten genereren, en geclassificeerd en al ter beschikking van het onderwijs zou moeten stellen. Een gedetailleerde 'inhoudsopgave' van zo'n bank is dan vereist om de bank voor de gebruiker toegankelijk te maken. Zo'n inhoudsopgave of domeinbeschrijving dient tevens als itemconstructiematrix. De gebruiksmogelijkheden van een domeinbeschrijving zijn echter veel ruimer. Er is onmiskenbaar behoefte aan een studie over de domeinen van het taalonderwijs. Recent heeft die behoefte zich doen gelden bij het opstellen van eindtermen (basisonderwijs en basisvorming) en bij onderzoek naar wenselijke doelstellingen zoals het onderzoek van Blok (1987) naar functionele taalvaardigheid. Het gaat om de behoefte doelen met elkaar in verband te brengen: om classificatie en taxonomie. In het verleden zijn er diverse pogingen gedaan, vooral op deelgebieden van het vak Nederlands en op het gebied van communicatieve competentie (tweede- en vreemdetaalonderwijs). Een studie die een theoretisch gefundeerd model oplevert dat orde kan brengen in de grote hoeveelheid voorstellen die reeds gedaan zijn, zal zeker lange tijd trend settend kunnen zijn.

In drie delen wordt getracht de weerbarstige klus te klaren het taalonderwijs in kaart te brengen. Het eerste deel (64 pagina's) bestaat naast een inleiding uit een hoofdstuk over enkele aspecten van de domeinbeschrijvingsproblematiek en een globale beschrijving van 'taalonderwijs in de praktijk'. Deel 2 (71 pagina's) bevat *het* model voor domeinbeschrijvingen. In vijf hoofdstukken worden vijf domeinen beschreven (in deze volgorde): de vier domeinen die te zamen de communicatieve competentie vormen (de linguïstische, de functionele, de tekstuele en de strategische competentie) en het sociaal-culturele domein. Het derde deel (ruim 40 pagina's) bevat twee hoofdstukken: een waarin 'de bruikbaarheid van het model als raamwerk voor domeinbeschrijvingen' aan de hand van enkele proefnemingen geëxploreerd wordt en een kort hoofdstuk 'conclusies en discussie'.

Het model

Kern van het proefschrift is deel 2, waarin het model voor domeinbeschrijvingen uiteengezet wordt. In het kort komt het op het volgende neer (zie Figuur 1).

De eerste vier domeinen ontleent JS aan 'recente

theorievorming omtrent communicatieve competentie'. Het vijfde domein heeft betrekking op een verzameling sociaal-culturele inhouden waarvan de doelen niet in termen van het vergroten van de taalvaardigheid kunnen worden geformuleerd.

De onderzoeker baseert de analyse van communicatieve competentie op het werk van Canale (1984) die een onderscheid maakt op basis van een taalkundige invalshoek (linguïstische, sociolinguïstische en tekstuele ('discourse') competentie), aangevuld met een strategische competentie, die ontleend wordt aan theorievorming en onderzoek in de toegepaste taalwetenschap.

Centraal staat het begrip communicatieve competentie. Aangenomen wordt dat er in communicatie sprake is van regelgestuurdheid en dat mensen die regels gebruiken op impliciete wijze in het concrete taalgebruik. JS vat die regelsystematieken op als onderliggende vaardigheden. Het algemene leerdoel in het talenonderwijs is volgens haar het vergroten of stimuleren van de communicatieve vaardigheden van leerlingen. Op methodische wijze tracht zij per onderscheiden domein tot specificaties van dat domein te komen. Ze onderscheidt drie 'hiërarchisch geordende wijzen van specificatie': begripsvalide specificaties, als wetenschappelijke theorievorming en onderzoek die kunnen leveren, specificatie op grond van 'common sense' en specificaties via curriculumonderzoek, te verkrijgen via analyse van de onderwijspraktijk.

Het domein van de *linguïstische competentie* (LC) bestaat uit de verzameling doelen die betrekking heeft op de beheersing van de code, van het taalsysteem als zodanig, bijvoorbeeld kennis en vaardigheden m.b.t. de uitspraak, de regels voor zins- en woordvorming, het lexicon. Het betekenisaspect bij LC beperkt zich tot de letterlijke, geïsoleerde, conventionele betekenis van woorden en zinnen, geïsoleerd van verdere contexten. Wie linguïstisch competent is, is in staat tot 'correct taalgebruik in overeenstemming met de taalbeschrijving zoals neergelegd in grammatica's en woordenboeken'. Sijtstra besluit (p. 70) dat linguïstische theorieën niet geschikt zijn om doelen m.b.t. de linguïstische competentie te specificeren. Ook zoeken naar specificaties waarover enige mate van consensus bestaat, levert niets op (p. 70). In minder dan twee pagina's beschrijft zij 'het traditionele onderwijs' dat volgens haar sterk gebaseerd is op het structuralisme in de taalkunde.

Het domein van de *functionele competentie* (FC) bevat de doelen die betrekking hebben op het situationeel bepaalde gebruik van taal: het gaat om *effectieve* toepassing van taal, passend in de gebruikssituatie door een passende keuze te maken - gezien de actuele situatie - uit het ter beschikking staande vorm- en betekenismateriaal. Hier treedt het sociale aspect van betekenis op de voorgrond. Implicatie is dat taalgebruikers de beschikking hebben over een arsenaal van taalmiddelen, waaruit een keuze kan worden gemaakt, specifiek toegesneden op de concrete situatie. De specificatiemogelijkheden van dit domein worden geleverd door de sociolinguïstiek, in het bijzonder het bekende taalgebruiksituatiemodel met componenten als code, 'topic', contact, 'setting', kanaal.

Het domein van de *tekstuele competentie* bevat de doelen m.b.t. de verwerking van teksten als geheel met de nadruk op het inhoudelijke aspect. "Een tekstueel competente taalgebruiker maakt onder meer gebruik van kennis en vaardigheden uit het linguïstische en functionele domein en is in staat tot zelfstandige en autonome verwerking van teksten van allerlei aard, gericht op het doorgronden van de inhoud daarvan"

	taalgebruiksmodus			
domein	luisteren	spreken	lezen	schrijven
linguïstische competentie				
functionele competentie				
tekstuele competentie				
strategische competentie				
sociaal-cultureel domein				

Figuur 1 Model voor domeinbeschrijving van het talenonderwijs volgens Sijtstra 1991

(p. 97). JS specificeert TC op basis van inzichten uit de psychologie van verwerkingsprocessen, waarbij zij aantekent dat over spreken erg weinig bekend is.

Het domein van de *strategische competentie* (SC) bevat de doelen die te maken hebben met het planmatig en bewust hanteren van de ter beschikking staande middelen van de communicatieve competentie. Strategische competentie behelst 'het inzicht in en de controle over de communicatie competentie' en is daarmee gedefinieerd als een metacomponent in het model. JS onderscheidt drie clusters van doelstellingen: strategieën in spontaan taalgebruik ('compensatie- en maximalisatiestrategieën') en in gestuurde taalverwerving (taalonderwijs) doelstellingen m.b.t. strategische taalbeschouwing (aspecten van metacommunicatieve processen en strategieën die een rol spelen als er zich een communicatief probleem voordoet (taalbeschouwing op linguïstisch, functioneel en tekstueel niveau) en doelstellingen m.b.t. 'de leerstrategische component', 'waarin processen van metacognitieve aard worden beschreven' (en die in het onderwijs bij het onderdeel studie- en informatievaardigheden aan bod komen). Het gaat dus om doelstellingen die te maken hebben met probleemoplossen bij communicatief handelen (metacommunicatief handelen) en bij het leren: metacognitieve strategieën spelen een rol als er zich een probleem voordoet in het eigen leren en/of denken.

Het *sociaal culturele domein* (SCD) bevat de doelen m.b.t. de beschouwing van taalverschijnselen en de daaraan gerelateerde aspecten van literaire, culturele en maatschappelijke aard (p. 135). Tot deze doelen behoren taalbeschouwing als culturele vorming ('Language Awareness'), het geïsoleerde grammatica-onderwijs, het literatuuronderwijs, en kennis van land en volk.

Toetsing van het model

De bruikbaarheid van het model wordt getoetst aan de curriculumvaliditeit van het model en wel aan de vakonderdelen schrijven (vreemdetalenonderwijs: einde voortgezet onderwijs; moedertaalonderwijs: einde basisonderwijs) en lezen (moedertaalonderwijs: einde basisonderwijs). JS gebruikt hiervoor toetsopgaven, die zij via een 'conceptuele analyse' classificeert met behulp van haar model. In principe wordt bij iedere opgave nagegaan welke noodzakelijke, niet triviale kennis en vaardigheid een leerling moet activeren teneinde een opgave tot een juiste oplossing te kunnen brengen. Deze procedure leidt tot het bepalen om welke competentie het gaat en tot, wat zij noemt, een uitgewerkte leerdoelstructuur.

Voor Schrijven in Moderne Vreemde Talen zijn *deelvaardigheidstoetsen* geanalyseerd. Een steekproef (105 opgaven) is door de onderzoekster geanalyseerd (vak Duits), daarna zijn 30 'representatieve items' (Frans) door twee onafhankelijk scorenden geclassificeerd met als resultaat: 'Slechts enkele items werden niet gelijk gescoord'. Ten slotte zijn twee *schrijftoetsen*, bestemd voor schoolonderzoeksdoeleinden, geanalyseerd. Het resultaat stemt de onderzoekster tevreden: enkele functionele opgaven bleken te behoren tot TC en enkele FC-opgaven zouden tot de strategische competentie behoren.

Voor Nederlands, niveau einde basisonderwijs, zijn 80 items uit *tekstbegriptoetsen* geanalyseerd. De items blijken voor het grootste deel te behoren tot het domein van de TC, behalve de items die behoren tot de directieve teksttypen (lezen van aanwijzingen e.d.). Deze behoren volgens JS tot FC, want 'ze zijn niet gericht op het doorgronden van de inhoud van de tekst'. Uit de analyse van 21 *schrijftaken* concludeert zij dat de meeste tot het domein van de FC behoren (p. 153) 'het kenmerkende van dat domein is de situationele inbedding' (...) 'achterliggende gedachte is dat de kenmerken van de concrete situatie bepalend zijn voor het actuele taalgebruik, i.v.m. de te produceren tekst'.

Later, in het slothoofdstuk Conclusies en discussie, doet JS verslag van nog een onderzoek naar de curriculumvaliditeit van het model. Dan gaat het om de vraag of de 'conceptie van taalvaardigheid empirisch gefundeerd' kan worden. Daartoe is een statistische analyse verricht op de gegevens van de eerste peiling van het onderwijsniveau (Nederlands, einde basisonderwijs). De scores op de tekstbegriptoetsen en schrijftaken (TC en FC) zijn als 'theoretische invoer' gebruikt. JS maakt uit de analyses op dat spelling, grammaticaliteit, interpunctie op LC laden en communicatieve effectiviteit, inhoud, stijl, en interactie op FC. Over TC is zij vaag. Een onderscheid tussen SC en FC/LC/TC heeft zij niet kunnen aantonen. Consequenties voor het ontwikkelde model en eerdere specificaties verbindt zij niet aan deze uitkomsten.

Evaluatie door Sijtstra

De onderzoeker is tevreden over het ontwikkelde model. Het voldoet volgens haar aan de eisen die zij stelde aan de modelontwikkeling en het model vervult tenminste drie pretenties.

Naast eisen die gelden voor iedere systematische beschrijving en classificatie (spaarzaamheid, communiceerbaarheid, praktische toepasbaarheid, werkbaarheid, eenduidigheid) stelt JS aan het te ontwikkelen model extra eisen, twee met betrekking tot de omvang, één met betrekking tot de structuur:

(1) Alle onderdelen van taalonderwijs moeten erin kunnen passen (spreken, luisteren, schrijven, lezen, taalbeschouwing en literatuur);

(2) Alle talen moeten erin zijn besloten: het onderwijs in het Nederlands (zowel moedertaalonderwijs als Nederlands als tweede taal) en het onderwijs in de moderne vreemde talen.

(3) Het model moet de doelstellingen in onderling aan elkaar gerelateerde systemen onder kunnen brengen, d.w.z., meer dan een opsomming van verschillende groepen van doelen bieden, dus het op een (...) 'overzichtelijke en samenhangende manier weergeven van doelen'. Dat kan alleen op basis van 'theoretische inzichten omtrent de aard van de doelen en inhouden van taalonderwijs, omtrent de aard van - kortweg - taalvaardigheid'. Daartoe wordt vanuit verschillende perspectieven, - de vakwetenschappen en algemene taalwetenschap, de psychologie, de onderwijskunde en de dagelijkse praktijk van het taalonderwijs- de problematiek van domeinbeschrijvingen bestudeerd.

Het resultaat van het onderzoek omgeeft JS met hoge pretenties: "Er is in kaart gebracht wat er met betrekking tot doel en inhoud kan worden ontrafeld gezien de huidige stand van zaken" (p. 11). "Het in dit kader ontwikkelde model ontstijgt aan de primaire context van toetsconstructie en is bruikbaar overal waar een systematisch vakinhoudelijk overzicht nodig is" (p. 13). "Het te beschrijven domein betreft de werkelijkheid zoals die in de klassepraktijk is aan te treffen" (p. 11).

Beoordeling

Geen van de pretenties wordt waargemaakt. Enerzijds is dat te wijten aan de opzet en uitvoering van het onderzoek, anderzijds aan de pretentieuze houding van de onderzoekster die het gehele proefschrift doortrekt. Zowel de theoretische onderbouwing van het model als het empirische onderzoek is onvoldoende om te kunnen spreken van een valide model.

De empirische toetsing. Het empirische gedeelte van de studie is zwak. De opzet is erg beperkt, de rapportage globaal en soms onnavolgbaar, de bespreking van de resultaten kritiek- en daarmee nogal zinloos.

Hoewel alleen toetsmateriaal is onderzocht, geen materiaal voor onderbouw basisonderwijs en voortgezet onderwijs is onderzocht, het aantal onderzochte vakonderdelen beperkt is, belangrijke vakonderdelen niet zijn onderzocht en minder belangrijke wel (schrijfvaardigheid bijvoorbeeld wel, dat in het vreemtalenonderwijs echter laag geprioriteerd is, en het hoger geprioriteerde spreekvaardigheid niet), Nederlands als tweede taal geheel buiten beschouwing is gebleven, besluit JS dat de verrichte proefnemingen een goede indicatie vormen voor de curriculumvaliditeit en de bruikbaarheid van het model. 'Op basis hiervan is het aannemelijk dat het model ook voor andere onderdelen van het reguliere curriculum als raamwerk voor domeinbeschrijvingen kan fungeren.' Dat zij in haar empirische onderzoek geen onderscheiding heeft kunnen aantonen tussen strategische competentie enerzijds en linguïstische, functionele en tekstuele competentie anderzijds, brengt haar er niet toe het model nog eens nader aan inspectie te onderwerpen. Dat de interpunctietoets blijkt te laden op een factor waarop ook spelling en grammaticaliteit laden, en het aspect stijl op een factor die JS definieert als Functionele Competentie, terwijl zij dat op basis van haar inhoudelijke analyses niet verwachtte, brengt de onderzoeker er niet toe haar inhoudelijke analyses nog eens onder de loep te nemen. Als zij constateert dat de classificatie van items uit schrijftoetsen maar zo weinig items van de oorspronkelijke indeling (het toetsconstructiemodel) af blijkt te wijken, interpreteert zij dat als een empirische onderbouwing van haar model en staat zij niet stil bij de toch voor de hand liggende mogelijkheid dat het toetsconstructiemodel blijkbaar óók een goed model is. Deze verzameling voorbeelden is uitbreidbaar en illustreert de weinig kwetsbare, reflectieve houding van de onderzoeker en het heilige geloof in het ontwikkelde model.

De theoretische fundering. Het proefschrift moet het hebben van de theoretische component. De theoretische fundering van het model is gelegd door Canadese onderzoekers (Canale & Swain, 1980). Erg opvallend in deze studie is de weinig schatplichtige houding die de onderzoeker tentoonspreidt aan het werk van Canale en Swain. JS zet deze fundering nergens helder uiteen: verspreid in het boek staan enkele citaten, die soms van commentaar worden voorzien. Een kritische analyse van die fundering ontbreekt; JS lijkt de onderscheidingen van Canale en Swain hier en daar wat aan te scherpen en verruimt de inhoud van de strategische competentie met leerstrategische, metacognitieve aspecten.

De kwaliteit van de literatuurstudie. De uitbreiding en herziening geschiedt op basis van een niet verantwoorde, beperkte selectie uit literatuur uit diverse disciplines. De basisonderscheidingen van Canale en Swain stammen uit onderscheidingen binnen de taalwetenschap: de linguïstiek, de sociolinguïstiek en de pragmalinguïstiek en tekstwetenschap ('discourse analysis'). De basisstructuur van het model van JS volgt die indeling: de zin (LC), de tekst (TC) en de tekst in gebruik (FC). Merkwaardig is dat zij voor de eerste twee domeinen een beroep doet op taalwetenschappelijke literatuur, maar de pragmalinguïstiek en 'discourse analysis' buiten beschouwing laat, zelfs de meer eenvoudige inleidingen als Leech, wat voor de onderbouwing en specificatie van het domein FC, maar ook voor de verhouding tussen TC en FC regelrecht een gemis is. Ook bijdragen uit de Nederlandse moedertaaldidactiek van Braet over de pragmatische dimensie in het leesonderwijs ontbreken ten onrechte. De belangrijke bijdrage van Blok aan de definiëring van functionele taalvaardigheid wordt niet behandeld, hoewel zijn proefschrift wel in de literatuurlijst prijkt. Een en ander heeft ertoe geleid dat de definiëring en specificatie van FC weinig hout snijdt en de relatie van FC tot TC niet uit de verf kan komen. Meer in het algemeen: de hier gepresenteerde literatuurstudie overtuigt niet: de fundering voor het model van deelcompetenties is smal. Daarbij valt op dat er veel literatuur niet is genoemd die een in dit domein ingevoerde lezer in één adem-tocht te binnen zou schieten, dat literatuur die wel genoemd wordt niet gebruikt wordt en dat waar wetenschappelijke literatuur voorhanden is soms gebruik wordt gemaakt van populair wetenschappelijke boekjes of schoolboeken. Een domein van literatuur waarbinnen ik veel mis, is theoretisch onderzoek dat is verricht t.b.v. het verkavelen van het doelstellingendomein van het moedertaal-, tweedetaal- en vreemdetaalonderwijs. Er zijn werkelijk lijsten aan te leggen van zaken die men daarin mist. Voor het moedertaalonderwijs bijvoorbeeld Moffett (1968), die de oude traditie van retorische analyse, die geheel in Sijtstra's studie ontbreekt, zeer belangrijk vernieuwde en wiens analyses nog steeds actueel zijn, en ten onzent Ten Brinke (1976) en Van Lint (1980). Ten behoeve van het vreemdetalenonderwijs hoeft men het niet in de besproken studie opgenomen standaardwerk van Stern (1983) maar door te bladeren om te zien dat belangrijke analytische werken op dit gebied zijn gemist.

Ook t.b.v. specificaties van de onderscheiden domeinen zijn storende leemtes waar te nemen. Men kan zich niet aan de indruk onttrekken dat de onderzoeker zich heel selectief, al dan niet zonder plan, door het domein heeft bewogen. Dat leidt tot onaanvaardbare reducties en vertekeningen. Zo haalt JS een schoolboek van Van Eemeren en Grootendorst aan, maar niet hun dissertatie waarop het schoolboek gebaseerd is, terwijl de classificatie van taalhandelingen in dat proefschrift een baanbrekend perspectief had kunnen bieden aan JS, die beweert dat al het werk van Searle en anderen m.b.t. taalhandelingen maar weinig heeft opgeleverd. Zo bestempelt JS een populair wetenschappelijke boekje van Dijkstra en Kempen als fundamenteel theoretisch en beweert ze dat er over spreekvaardigheid slechts weinig theoretische inzichten zijn ontwikkeld, voorbijgaand aan het standaardwerk van Levelt en alle literatuur binnen de wetenschappelijke discipline van 'discourse analysis'. Zo beweert JS over het vakonderdeel taalbeschouwing dat er weinig analyses voorhanden zijn, daarmee veel werk van de onderwijstaalkundigen Van Dort, Van Calcar, Hulshof en Sturm negerend. Zo wordt voor het vakonderdeel lezen naast diverse handboeken over het leesproces node de uitwerking van de handelingstheoretische invalshoek van Bol gemist, die het mogelijk had gemaakt dat de onderzoekster meer oog had gehad voor een functionele, communicatieve invalshoek van het lezen. En zo is, ten slotte, de onderbouwing van het sociaal-

culturele domein heel mager. Voor het vakonderdeel literatuur ontbreekt al het werk van de Nijmeegse literatuurdidacticus De Moor en zijn medewerkers en belangrijke bijdragen die Purves leverde voor de verkaveling van het doelstellingendomein.

De theoretische onderscheidingen. Verspreid door het boek behandelt JS grensgevallen: vakonderdelen of items die in het ene dan wel het andere domein geplaatst dienen te worden. In maar weinig gevallen overtuigen die besprekingen. Zo is het zeer betwijfelbaar of hardoplezen alleen een beroep doet op linguïstische competentie en het op juiste toon lezen niet vereist dat betekenistoekenning onder meer berust op de directe context van de te lezen zin. Zo is het ook betwijfelbaar waartoe de doelstellingen van ontleedonderwijs behoren. Volgens JS niet tot LC (p. 77), want de doelen binnen LC hebben betrekking op de 'vaardige beheersing van aspecten van het taalsysteem' en ontleedonderwijs zou niet met dat oogmerk gegeven worden. Het is echter wel zeker dat veel ontleedonderwijs wordt gegeven, al dan niet terecht, met als legitimatie (vreemde) taalvaardigheidsvergroting.

Het onderscheid tussen TC/FC en SC dat JS maakt, is twijfelachtig. Tekstbegrip rekent zij tot TC, maar als leesactiviteiten bewuste aandacht vereisen, zoals studerend lezen, worden ze ondergebracht bij strategische competentie. Ook als het lezen (of het beantwoorden van vragen over een tekst) een beroep doet op kennis van de wereld, wordt de activiteit niet meer tot TC gerekend maar tot SC: een vrijwel onhoudbare stellingname. Het onderbrengen van 'monitoring'-activiteiten bij SC in plaats van bij FC roept ook veel vraagtekens op. De FC verzorgt de analyse van parametersetting van de communicatieve situatie: aanpassing van de betekenis en vorm aan de situatie kan daardoor plaats vinden. Maar de FC draagt ook bij aan de continue monitoring van de taalgebruikssituatie en de evaluatie van de 'passing' om bij een mogelijk optredend betekenisprobleem (zowel referentieel als relationeel) adequaat te kunnen reageren. Zeker als het gaat om maximalisatie- en compensatiestrategieën ligt een onderbrenging in de FC veel meer voor de hand dan in de SC.

De uitbreiding van de SC met metacognitieve strategieën, - kennis van het eigen denk- en leerproces, die wordt ingezet om sturing te geven aan en controle uit te oefenen over het denken en leren - is interessant. In het taalonderwijs wordt, op doelstellingenniveau, steeds meer aandacht besteed aan het leren leren: hoe vergroot je je taalverwervingscapaciteit? Hoe vergroot je je communicatieve competentie? Het spreekt voor zich dat deze taalleercomponent gebruik maakt van de planning-, monitor,- en evaluatiecomponent van communicatief handelen. Maar het gaat om andere leeractiviteiten dan die JS als voorbeelden noemt: studielessen, -vaardigheden, informatievaardigheden. Dit zijn te breed gedefinieerde activiteiten voor het *taal*onderwijs. Het zijn eerder leeractiviteiten waarbij taalvaardigheid een rol speelt, dan taalleeractiviteiten. Het doelstellingendomein van het taalonderwijs wordt zo ongelukkigerwijs veel te breed: alle leren op school is dan taalonderwijs.

Een centraal begrip in deze studie, taalvaardigheid, krijgt maar heel weinig aandacht. Zo gaat JS er vanuit dat de vier onderscheiden modi (spreken, luisteren, schrijven en lezen) recht van bestaan hebben: volgens JS is er 'enige empirische basis' voor deze onderscheidingen, zonder te verwijzen naar literatuur (p. 28). Over deze kwestie is echter veel empirisch onderzoek voorhanden waaronder onderzoek dat deze stelling in twijfel trekt. Zij refereert daaraan niet. Keer op keer blijkt in onderzoek weer dat het bijvoorbeeld nog maar de vraag is in hoeverre er zelfs sprake is van schrijf- en spreekvaardigheid, zo zwak hangen prestaties op taken binnen hetzelfde domein samen. De studie bevat meer van dit soort boude beweringen en aannames die niet of spaarzaam ondersteund worden door argumentatie en/of empirische gegevens. Zo gaat JS er vanuit dat leesvaardigheidsprocessen in de vreemde taal op dezelfde wijze verlopen als in de moedertaal. Ze stelt bijvoorbeeld ook, zonder verdere consequenties te trekken, dat elk type tekst een bepaalde manier van lezen vergt.

Het is het streven naar alomvattendheid dat haar opbreekt, ook waar te nemen bij de behandeling van het buitenbeentje der competenties. Het sociaal-culturele domein is zo stiefmoederlijk behandeld, dat dit domein, in deze een vergaarbak allerhande, veel beter buiten beschouwing had kunnen blijven. Het belang dat men in de onderwijspraktijk aan dit domein hecht - in de bovenbouw van het h.a.v.o. en v.w.o. besteedt men gemiddeld 35% van de onderwijstijd aan

literatuuronderwijs - verdient meer en betere aandacht. Als er toch voor breedte moest worden gekozen, zou meer aandacht aan spreek- en luistervaardigheid en het interactieve karakter van communicatie beter geweest zijn. Die aandacht had niet misstaan omdat het een veel te belangrijk terrein is van taalonderwijs en de taalonderwijspraktijk. Die aandacht had een bijdrage kunnen leveren aan een betere definiëring van de wat JS noemt Functionele Competentie en een betere afbakening daarvan tot Tekstuele en Strategische Competentie.

Ik vind het jammer dat dit proefschrift zo weinig kan bijdragen aan de studie over en de praktijk van het taalonderwijs. Tevoren had het onderzoek veel sterker beperkt moeten worden. In principe is het mogelijk dat een onderzoeker tot diezelfde conclusie komt in het slothoofdstuk, als de opbrengst en het onderzoeksproces kritisch tegen het licht wordt gehouden. In dit proefschrift is dat niet het geval. Een kritische ondertoon geldt alleen andermans werk en het in dit onderzoek ontwikkelde model wordt door de onderzoeker omgeven door loftuitingen en pretenties. De bijdrage die het proefschrift had kunnen leveren aan het onderzoek van doelstellingen door een reflectieve, zich kwetsbaar opstellende onderzoeker die wikt en weegt, levert dit proefschrift helaas niet.

G. Rijlaarsdam

Literatuur

Brinke, J. S. ten (1976). *The complete mothertongue curriculum*. Groningen: Wolters-Noordhoff-Longman.

Blok, H. (1987). *Taal voor alledag*. Dissertatie Universiteit van Amsterdam. 's-Gravenhage: SVO.

Bol, E. (1982). *Leespsychologie*. Groningen: Wolters-Noordhoff.

Braet, A. (1979) De pragmatische dimensie van taalbeheersing 1: van tekstbehandeling naar communicatie-analyse. *Levende Talen, 310*, 68-76.

Canale, M. (1984). A communicative approach to language proficiency assessment in a minority setting. In Ch. Rivera (Ed.), *Communicative competence*. Approaches to language proficiency assessment, research and application. Clevedon: Multilingual Matters.

Canale, M, & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1, 1-47.

Dijkstra, T., & Kempen, G. (1984). *Taal in uitvoering*. Inleiding tot de psycholinguïstiek. Groningen: Wolters-Noordhoff.

Eemeren, F. van, Grootendorst, R., & Kruiger, T. (1983). *Argumentatieleer*. Deel 1. Het analyseren van een betoog. Groningen: Wolters-Noordhoff.

Eemeren, F. H. van, & Grootendorst, R. (1982). *Regels voor redelijke discussies*. Dissertatie Universiteit van Amsterdam. Dordrecht: Foris.

Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.

Lint, P. van (1980). *Vormen van taalverkeer*. Een elementaire inleiding in de leer van de verstandhoudingsmiddelen: de mondelinge vormen. Groningen: Xeno.

Moffett, J. (1968). *Teaching the universe of discourse*. Boston: Houghton Miffling Company.

Purves, A. C. (1971). Evaluation of learning in literature. In B. S. Bloom, et al., *Handbook on formative and summative evaluation of student learning* (pp. 697-766). New York.

Stern, H. H. (1983). *Fundamental Concepts of language teaching*. Oxford: Oxford University Press.

Mededelingen

PAO-prijs 1993

De Stichting Nederlands Instituut voor Postacademisch Onderwijs (PAO) stelt jaarlijks een prijs van *f* 3.000,- ter beschikking voor het beste overzichtsartikel op het gebied van de kinder- en jeugdpsychologie en daaraan verwante vakgebieden als (ortho) pedagogiek, onderwijskunde en kinder- en jeugdpsychiatrie. Het artikel dient actueel wetenschappelijk onderzoek op deze gebieden toegankelijk te maken voor hen die in de praktijk werkzaam zijn, en dient bij voorkeur een nieuw perspectief te bieden.

De toekenning van de PAO-prijs gebeurt door een commissie, samengesteld uit vertegenwoordigers van de Sector Jeugd van het Nederlands Instituut van Psychologen en de Werkgemeenschap Ontwikkelingspsychologie van de Stichting Psychon, onder voorzitterschap van prof.dr. W. Koops van de Vrije Universiteit.

Bij de toekenning worden de volgende maatstaven aangelegd:

1. Het artikel moet in 1993 zijn c.q. worden gepubliceerd in een geregistreerde wetenschappelijke publikatie (tijdschrift of boek met resp. ISSN- of ISBN-nummer).
   Maximum omvang: 7500 woorden.
2. Het artikel moet betrekking hebben op een thema dat relevant is voor in de praktijk werkzame psychologen, pedagogen, psychiaters.
3. De hoofdzaken uit de onderzoeksliteratuur van een aangegeven tijdvak moeten zijn beschreven.
4. De wetenschappelijke betekenis van de besproken literatuur dient duidelijk te zijn aangegeven.
5. De implicaties van de besproken literatuur voor de beroepspraktijk dienen duidelijk te zijn aangegeven.

Inzendingen, voor de prijs van 1993 in te dienen tot 1 januari 1994, kunnen gericht worden aan: Secretariaat Commissie PAO-prijs, Postbus 325, 2300 AH Leiden, tel. 071-143142, fax 071-140145.

N.B. Zij die bekend zijn met een of meer artikelen die bij uitstek lijken te voldoen aan de aangegeven criteria worden uitgenodigd om de desbetreffende auteur(s) te wijzen op de mogelijkheid om het artikel in te zenden voor de PAO-prijs.

Landelijk Congres Vrouwen/ Genderstudies

Het eerste interdisciplinaire Landelijk Congres Vrouwen/Genderstudies vindt plaats op 28 en 29 oktober 1993 in het hoofdgebouw van de Vrije Universiteit te Amsterdam. Hier presenteren en bespreken onderzoekers de resultaten van hun onderzoek en wordt de balans opgemaakt: wat heeft twintig jaar vrouwenstudies opgeleverd, wat zijn de lacunes en wat is de toekomst van vrouwen/genderstudies in Nederland. Naast de paperpresentaties binnen de drie thema's: 'generaties en overdracht', 'macht en marginaliteit' en 'materie en constructie', is er een plenair programma. Dit omvat lezingen, inleidingen, een prijsuitreiking en de forumdiscussie over de 'Pretenties en potenties van vrouwenstudies'.

Aanmelding en informatie: SISWO/Instituut voor Maatschappijwetenschappen, Plantage Muidergracht 4, 1018 TV Amsterdam, tel. 020-5270600 (K. Corver, D. van der Heiden).

Weer samen naar school

Onder auspiciën van het NIP en de NVO wordt op 19 november 1993 in De Eenhoorn te Amersfoort een conferentie over dit thema gehouden.

Voor nadere informatie kunt u zich wenden tot het NVO Secretariaat, Korte Elisabethstraat 11, 3511 JG Utrecht, tel. 030-322407.

Studiedag over Pedagogische Preventie

Op 12 november 1993 organiseert de Vakgroep Pedagogiek van de Universiteit Utrecht de studiedag 'Van kindmodel naar modelkind? Theorie praktijk van preventie van psychosociale problemen bij kinderen en jeugdigen'. De ochtend bestaat uit een serie theoretische bijdragen over preventie, risicofactoren en beschermende factoren. In de middagsessies komen een aantal concrete preventie-activiteiten aan bod over een zestal thema's.

Nadere informatie is verkrijgbaar bij de Vakgroep Pedagogiek, mevrouw drs. I. Groenestege,

tel. 030-535779 (inhoudelijk) en mevrouw S. Eilerts, tel. 030-534601 (organisatorisch).

Internationaal Dyslexie Congres

Op 27 mei 1994 wordt te Maastricht een internationaal dyslexie congres gehouden.

Voor nadere inlichtingen: Stichting Dyslexie Limburg, tel. 04492-3528.

Internationaal Congres Ontwikkelingspsychologie - 1994

Thirteenth Biennial Meetings of ISSBD

Op 28 juni tot en met 2 juli 1994 vindt in Amsterdam het dertiende tweejaarlijkse congres plaats van de ISSBD, de International Society for the Study of Behavioural Development. Dit is een organisatie van ruim 1200 wetenschappers uit meer dan 50 landen die vanuit verschillende disciplines als psychologie, pedagogiek en psychiatrie betrokken zijn bij de studie van menselijke ontwikkeling in alle levensfasen.

Belangrijkste doel van het congres is om onderzoekers de gelegenheid te geven hun werk te presenteren aan en bediscussiëren met collega's uit andere landen en andere vakgebieden.

Daarnaast is het ook een uitnemende gelegenheid voor in de praktijk werkzame psychologen, pedagogen en psychiaters om kennis te nemen van de huidige stand van zaken in het wetenschappelijk onderzoek op voor hen relevante gebieden.

Een internationale programmacommissie onder voorzitterschap van prof.dr. W. Koops (VU) is verantwoordelijk voor het wetenschappelijk programma dat zal bestaan uit lezingen door uitgenodigde sprekers, symposia, posterpresentaties en posterworkshops, en speciale evenementen.

Tot 30 september 1993 is het mogelijk om voorstellen voor posters of poster workshops in te dienen. De daartoe bestemde formulieren zijn gevoegd bij de congresfolder die verkrijgbaar is bij: PAO, Postbus 325, 2300 AH Leiden, tel. 071-143142, fax 071-140145.

Inhoud andere tijdschriften

Comenius
13e jaargang, nr. 1, 1993

Introductie op het thema: Praktijken en theorieën van multiculturaliteit, door W. Wardekker, S. Miedema en H. Baartman
Cultuurverschillen in een multiculturele samenleving, door L. Eldering
De koppigheidsfase. Een reconstructie van een cultureel gedefinieerd opvoedingsprobleem, door J. van der Zwaard
Het belang van een cultuurbeleid in het onderwijs, door T. Pels
Pluralisme in (re)actie, door H. Procee
Op spokenjacht. Lessen over inter-etnisch samenleven en discriminatie in het voortgezet onderwijs, door Y. Leeman

Tijdschrift voor Onderwijsresearch
17e jaargang, nr. 6, 1992

De multiniveau-benadering toegepast op leesvorderingen, door M. W. Vooijs, L. J. Th. van der Kamp, C. M. Koolstra en T. H. A. van der Voort
Een vergelijking van twee schatters voor de intraklasse correlatie coëfficiënt: een simulatie in de context van het onderzoek naar de effectiviteit van scholen, door M. H. Eiting en H. Blok
De grootte van het schooleffect in het basisonderwijs; een analyse op basis van vijf jaar Eindtoets Basisonderwijs, door H. Blok
Psychometrische kwaliteiten van meetinstrumenten van het onderwijsaanbod, door E. Harskamp en C. Suhre

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 5, 1993

Leven (z)onder vrienden, door J. D. van der Ploeg
Aantekeningen bij leven (z)onder vrienden, door J. M. A. M. Janssens
Het oog van de meester. Enige orthopedagogische opmerkingen over het kijken naar leerlingen in het gespecialiseerd onderwijs, door W. ter Horst
Het lezen van pseudowoorden en bestaande woorden. Deel II: Decodeerroutes en decodeerproblemen, door K. P. van den Bos en A. J. M. Scheepstra

Ontvangen boeken

Baartman, H. E. M., *Opvoeden met alle geweld. Hardnekkige gewoontes en hardhandige ouders.* Uitgeverij SWP, Utrecht, 1993, *f* 39,50.

Bruin-de Boer, A. de, & Greef, E. de, *Hoogintelligente kinderen in het basisonderwijs.* De Ruiter, Gorinchem, 1993, *f* 47,50.

Cornelissen, F., & Dekkers, P., *Schoolwerkplanontwikkeling.* Moduleboek. Wolters-Noordhoff, Groningen, 1993, *f* 27,50.

Dijkstra, A., & Kempen, G., *Taalpsychologie.* Wolters-Noordhoff, Groningen, 1993, *f* 39,50.

Hamstra-Bletz, L, *Het kinderhandschrift. Ontwikkeling en beoordeling* (dissertatie). Universiteit Leiden, 1993.

Heetveld, G., *Individualiseren en differentiëren.* Moduleboek. Wolters-Noordhoff, Groningen, 1993, *f* 27,50.

Huisman, A., *Pedagogisch beleid voor kindercentrum en gastouderbureau.* Uitgeverij SWP, Utrecht, 1993, *f* 26,50.

Joolingen, W. van, *Understanding and facilitating discovery learning in computer-based simulation environments* (dissertatie). Eindhoven University, 1993.

Koenraadt, F. (red.), *Forensische jeugdpsychologie. Rapportage en behandeling in het familie- en jeugdrecht.* Gouda Quint, Arnhem, 1993, *f* 52,–.

Moust, J., *De rol van tutoren in probleemgestuurd onderwijs. Contrasten tussen student- en docenttutoren* (dissertatie). Universitaire Pers, Maastricht, 1993.

Plass, S., & Mulder, A. (red.), *Suïcide bij jongeren.* Bohn Stafleu Van Loghum, Houten, 1993, *f* 39,50.

Raymakers, L. P. T. (eindred.), *Zorg voor pedagogische kwaliteit.* Dekker & Van de Vegt, Assen, 1993, *f* 29,50.

Tavecchio, L. W. C., & Hoogeweij, J., *Oefenboek Statistiek – Methoden en Technieken.* Wolters-Noordhoff, Groningen, 1993, *f* 49,50.

Veugelers, W., *Pedagogische opdracht en arbeid. Doelstellingen van docenten in het Algemeen Voortgezet Onderwijs en het Beroepsonderwijs* (dissertatie). Thesis Uitgeverij, Amsterdam, 1993, *f* 35,–.

Wielemans, W., *Voorbij het individu.* Uitgeverij Garant, Apeldoorn/Leuven, 1993, *f* 46,75.

ZORG VOOR PEDAGOGISCHE KWALITEIT

Drs. L.P.T. Raijmakers (eindredactie)

1993. 148 blz. *f* 29,50
ISBN 90 255 0082 X
NVO-Publikaties, 4

Een belangrijk boek voor een ieder die betrokken is bij de zorg voor kwaliteit in de hulpverlening of in het onderwijs; voor diegenen die te maken hebben met het kwaliteitsbeleid van de overheid, maar die vanuit de eigen instelling ook een eigen kwaliteitsbeleid dienen te ontwikkelen.

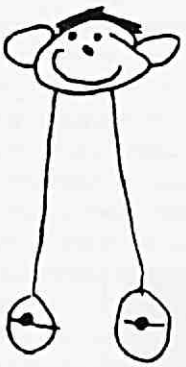
Verkrijgbaar in de boekhandel of rechtstreeks bij de uitgever

Uitgeverij Dekker & van de Vegt
Tel. 05920-40995

Besteladres:
Uitgeverij VAN GORCUM, Postbus 43, 9400 AA Assen, Tel. 05920-46846/Fax 05920-72064

Swets & Zeitlinger Publishers

Heereweg 347, 2161 CA Lisse



Tekenen – een natuurlijke bezigheid

De psychologie van kindertekeningen

Glyn V. Thomas en Angèle M.J. Silk

ISBN 90 265 1318 6 180 PAG. *f* 29,50

Het meest recente en uitgebreide overzicht van psychologische theorieën over kindertekeningen. Waarom tekent een kind? Hoe tekent een kind? Wat zegt een tekening over ontwikkeling, niveau, begaafdheid, originaliteit, emoties, expressiviteit? Vanuit de ontwikkelingspsychologie, de klinische psychologie, de artistieke benadering en de cognitieve psychologie kijkt men naar verschillende aspecten. Recent onderzoek is vooral gericht op de analyse van het tekenproces en op het vaststellen van de informatie die kinderen in hun tekeningen verwerken.
Vertaling van *An Introduction to the Psychology of Children's Drawings*.
De auteurs zijn als psycholoog verbonden aan de Universiteit van Birmingham (Engeland).

Opvoeding en opvoedingsproblemen in het gezinsleven

P.M. Schoorl

ISBN 90 265 1265 1 184 PAG. *f* 45,—

In elk gezin ontstaan vanzelfsprekende handelingspatronen, waarvan de gezinsleden zich niet altijd bewust zijn. Dit boek geeft een handleiding voor het blootleggen van deze impliciete gezinsregels in probleemsituaties, als noodzakelijke voorbereiding op de hulpverlening. De basis is een theoretische analyse van de regels in het gezinsleven en hun invloed op het functioneren van het gezin. Het resultaat is een praktisch en overzichtelijk boek over gezinsopvoeding, dat door de heldere en systematische opbouw ook zeer geschikt is voor opleidingen.

Verkrijgbaar in de boekhandel of bel: 02521 - 35111

15 oct 6441

Inhoudsopgave

Pedagogische Studiën Jaargang 70, nummer 4/1993

L. Mulder
De invloed van het advies op de schoolloopbaan van de OVB-doelgroepleerlingen in het voortgezet onderwijs 242

G. Driessen
Binnen en buiten Onderwijsvoorrangsgebieden. Sociaal-etnische achtergronden en onderwijspositie van leerlingen in het voortgezet onderwijs 252

C. Verhaak
Samenwerking in Onderwijsvoorrangsgebieden 263

F. K. Kieviet
Visitaties lerarenopleidingen nader bekeken 274

H. W. C. H. van Amelsvoort, R. J. Bosker en T. F. Brandsma
Rendementsbekostiging in de bovenbouw van HAVO en VWO; overwegingen en reacties 283

A. M. A. Walraven, P. Reitsma en E. J. Kappers
Instructie in strategieën voor begrijpend lezen aan zwakke lezers 298

Kroniek
J. H. Slavenburg
Voorschoolse stimuleringsprojecten 309

Boekbespreking 316

Mededelingen 318

BIBLIOTHEEK DER
RIJKSUNIVERSITEIT
UTRECHT

De invloed van het advies op de schoolloopbaan van de OVB-doelgroepleerlingen in het voortgezet onderwijs*

L. Mulder**

Samenvatting

Het advies voor voortgezet onderwijs is voor het grootste deel gebaseerd op de leerprestaties van leerlingen. De relatie tussen schoolprestaties en advies is echter niet geheel eenduidig. De afgelopen jaren hebben diverse onderzoekers melding gemaakt van hun bevinding dat er bij de adviezen aan allochtone leerlingen mogelijk lagere eisen worden gesteld dan aan autochtone leerlingen. Dit roept de vraag op welke gevolgen dat heeft voor de verdere loopbaan in het voortgezet onderwijs. In dit artikel wordt die vraag nader onderzocht op basis van gegevens uit de landelijke evaluatie van het onderwijsvoorrangsbeleid. Het blijkt dat allochtone leerlingen inderdaad relatief hoge adviezen krijgen, maar alleen in vergelijking met autochtone leerlingen met vergelijkbare sociale achtergrond. De relatief hoge adviezen hebben geen nadelig effect op de onderwijspositie na twee jaar voortgezet onderwijs. Dat de onderwijspositie van de allochtone doelgroepleerlingen in het tweede leerjaar ongunstiger is dan die van de overige leerlingen, is voornamelijk een gevolg van het feit dat allochtone leerlingen relatief vaak van het advies afwijken en voor een hoger schooltype kiezen dan is geadviseerd.

Inleiding

Het uiteindelijke doel van het onderwijsvoorrangsbeleid (OVB) is, via een verbetering van de taal- en rekenprestaties in het basisonderwijs, de opleidings- en beroepsmogelijkheden van met name allochtone en autochtone arbeiderskinderen te verruimen. Het effect van het OVB zal dus op den duur tot uitdrukking moeten komen in een grotere deelname van deze leerlingen aan hogere vormen van voortgezet onderwijs en het bereiken van hogere onderwijseindniveaus. In de OVB-evaluatie wordt daarom ruimschoots aandacht besteed aan de vraag hoe de onderwijsachterstanden van de doelgroepleerlingen zich binnen het voortgezet onderwijs ontwikkelen. Om daar uitspraken over te kunnen doen, worden ruim 5000 leerlingen die in 1988 bij de eerste meting van de OVB-evaluatie in het basisonderwijs in groep 8 zijn getoetst, ook in het voortgezet onderwijs verder gevolgd. Over deze leerlingen worden jaarlijks gegevens opgevraagd over schooltype, zittenblijven, op- en afstroom en de behaalde rapportcijfers. Op dit moment zijn de gegevens over de eerste twee leerjaren bekend.

De overgang van basis- naar voortgezet onderwijs is een belangrijk en grotendeels bepalend moment in de totale gang van leerlingen door het onderwijs. Met de keuze voor het type voortgezet onderwijs worden immers de beroepsmogelijkheden en daarmee de maatschappelijke kansen al in grote lijnen vastgelegd. Hoe de keuze uitvalt, hangt af van allerlei factoren. Een van de belangrijkste factoren is het advies van de basisschool.

Uit diverse onderzoeken is gebleken dat het advies dat aan allochtone leerlingen wordt gegeven op een andere manier tot stand komt dan het advies aan autochtone leerlingen (De Jong 1987; Driessen, 1991; Mulder, 1991). De onderzoeksresultaten leidden vrijwel steeds tot de conclusie, of in ieder geval tot het sterke ver-

* De in dit artikel gepresenteerde gegevens zijn door het ITS te Nijmegen en het RION te Groningen verzameld in het kader van de landelijke evaluatie van het onderwijsvoorrangsbeleid. De landelijke evaluatie is een onderzoeksprogramma van de Stichting voor Onderzoek van het onderwijs (SVO).

** Met dank aan G. Driessen voor zijn commentaar op een eerdere versie van dit artikel.

moeden, dat allochtone leerlingen bij de advisering positief worden gediscrimineerd. Zij zouden een hoger advies krijgen dan autochtone leerlingen met vergelijkbare prestaties. Wanneer dat inderdaad zo is, rijst de vraag welke gevolgen dat heeft voor de verdere schoolloopbaan in het voortgezet onderwijs. Blijven allochtone leerlingen door het relatief hoge advies vaker zitten of gaan ze vaker naar een ander schooltype? Of blijkt het advies een goede voorspeller te zijn voor de verdere schoolloopbaan in de eerste jaren van het voortgezet onderwijs?

In dit artikel proberen we een antwoord te vinden op deze vragen. We doen dat op basis van gegevens van de leerlingen die in het kader van de OVB-evaluatie in het voortgezet onderwijs worden gevolgd.

1 Steekproef en variabelen

Van de 13.000 leerlingen die bij de eerste meting van de OVB-evaluatie in groep 8 zijn getoetst is ongeveer de helft geselecteerd voor verder onderzoek in het voortgezet onderwijs. Bij de selectie van die leerlingen is uitgegaan van de steekproefopzet van de OVB-cohort-onderzoeken. Uit een referentiesteekproef van de cohort-onderzoeken zijn op toevalsbasis 140 basisscholen getrokken. De leerlingen op die scholen kunnen representatief worden geacht voor alle leerlingen die in 1989 het basisonderwijs hebben verlaten. Deze steekproef is aangevuld met 200 basisscholen waar relatief veel leerlingen uit de OVB-doelgroepen op zitten. Het totale aantal geselecteerde leerlingen kwam hiermee op circa 6000. Het was de bedoeling dat deze 6000 leerlingen in het voortgezet onderwijs zouden worden gevolgd. Door non-respons van basisscholen en scholen voor voortgezet onderwijs is dit aantal niet gehaald. In het eerste leerjaar van het voortgezet onderwijs zijn gegevens verzameld over 5322 leerlingen en in het tweede leerjaar over 4965 leerlingen. Uitvalsanalyses hebben uitgewezen dat de non-respons geen nadelige gevolgen heeft gehad voor de representativiteit van de referentiesteekproef (Mulder, 1991; Mulder & Pijl, 1992).

In het najaar van 1989 zijn de basisscholen die bij de OVB-evaluatie zijn betrokken benaderd met het verzoek op te geven welk advies voor voortgezet onderwijs de leerlingen die het jaar daarvoor in groep 8 waren getoetst, hadden gekregen. Ook is gevraagd bij welke school voor voortgezet onderwijs de leerlingen zich hadden aangemeld.

Volgens opgave van de basisscholen zaten de geselecteerde leerlingen verspreid over meer dan 1100 verschillende scholen voor voortgezet onderwijs. Aan deze scholen is in januari 1990 en in januari 1991 schriftelijk gevraagd of de leerlingen (nog steeds) op de school zaten en zo ja, op welke schooltype en in welk leerjaar. Op basis hiervan en van de gegevens die in het basisonderwijs al waren verzameld, wordt in dit artikel de relatie beschreven tussen de taal- en rekenprestaties in groep 8, het advies, de feitelijke schoolkeuze en de onderwijspositie na twee jaar voortgezet onderwijs.

2 Resultaten

2.1 Prestaties en adviezen in het basisonderwijs

Op het moment dat leerlingen van het basisonderwijs naar het voortgezet onderwijs gaan, hebben zij al een groot deel van hun schoolloopbaan achter de rug. Voor de ene leerling is die meer succesvol verlopen dan voor de andere. De mate van succes komt voornamelijk tot uiting in de schoolprestaties en in het aantal keren doubleren. Uit de resultaten van de eerste meting van de OVB-evaluatie blijkt dat leerlingen uit de OVB-doelgroepen op beide aspecten ongunstig uit de bus komen. Zo blijven Nederlandse en allochtone doelgroepleerlingen veel vaker zitten dan leerlingen die niet tot de OVB-doelgroepen behoren. Uitgedrukt in percentages is de verdeling over de OVB-categorieën als volgt: van de 1.00-leerlingen blijft 10 procent in het basisonderwijs wel eens zitten, van de 1.25-leerlingen 17 procent en van de 1.90-leerlingen 27 procent. Ondanks het feit dat veel leerlingen uit de OVB-doelgroepen een jaar langer over de leerstof doen zijn hun taal- en rekenprestaties aan het eind van het basisonderwijs nog steeds lager dan die van de overige leerlingen. Hoe groot de achterstand is, staat in Tabel 1. De scores in de tabel zijn zogenaamde T-scores: standaardnormaalscores met een ge-

Tabel 1

T-scores voor taal en rekenen in groep 8, naar wegingsfactor. Gemiddelden (X), standaarddeviaties (sd) en aantallen (N)

	TAAL			REKENEN		
wegingsfactor	*X*	*sd*	*N*	*X*	*sd*	*N*
1.00	52.3	9.3	2509	52.3	9.6	2521
1.25	48.5	9.5	1942	48.3	9.5	1927
1.90	39.3	9.4	2544	42.1	9.6	2542

middelde van 50 en een standaarddeviatie van 10. De T-scores zijn berekend op basis van de taal- en rekenscores van de leerlingen uit de referentiesteekproef. De T-scores geven de afwijking ten opzichte van het gemiddelde van de totale groep weer. Hierdoor kunnen scores van verschillende groepen leerlingen en van verschillende toetsen met elkaar worden vergeleken. De aantallen in de tabel hebben betrekking op alle leerlingen die in 1988 in groep 8 zijn getoetst. Dat zijn er meer dan het aantal dat in het voortgezet onderwijs wordt gevolgd.

De achterstand van de OVB-doelgroepleerlingen komt in Tabel 1 duidelijk naar voren. Zowel bij taal als bij rekenen blijven de 1.25- en 1.90-leerlingen onder het gemiddelde van de referentiesteekproef. Vooral de achterstand van de 1.90-leerlingen is bijzonder groot. Wat taalprestaties betreft ligt deze groep meer dan een standaarddeviatie onder het landelijke gemiddelde. De rekenprestaties zijn beter, maar ook deze liggen nog ver onder de 50. In vergelijking met de prestaties van de allochtone doelgroepleerlingen valt de achterstand van de 1.25-leerlingen mee. Zowel bij taal als bij rekenen zijn hun scores ongeveer eenzesde standaarddeviatie lager dan het landelijke gemiddelde. De gemiddelde scores laten echter niet zien dat er binnen de 1.25-leerlingen een groep leerlingen is te onderscheiden die vrijwel geen achterstand heeft. Het gaat daarbij om leerlingen van wie één van beide ouders een opleiding hoger dan LBO heeft of een beroepsniveau hoger dan geschoolde handarbeid. De taal- en rekenscores van deze leerlingen wijken nauwelijks van die van de 1.00-leerlingen af en daardoor verhullen ze de achterstand van de echte probleemgroep: de leerlingen van wie beide ouders een laag opleidings- en beroepsniveau hebben. De gemiddelde T-score op taal en rekenen van de 'echte' 1.25-leerlingen ligt rond de 46 (Tesser, Mulder & Van der Werf, 1991).

Het spreekt vanzelf dat leerlingen die met een achterstand het basisonderwijs afsluiten, niet voor alle schooltypen in het voortgezet onderwijs geschikt zijn. Het is de taak van de basisschool om een advies uit te brengen over het schooltype dat het best bij de betreffende leerling past. Gezien de taal- en rekenprestaties van de leerlingen uit de OVB-doelgroepen is het niet verwonderlijk dat zij over het algemeen lage adviezen krijgen. In Tabel 2 is te zien hoe de verdeling van de adviezen er per wegingsfactor uitziet.

Tabel 2

Adviezen uitgesplitst naar wegingsfactor

	1.00	1.25	1.90
IBO	.1	2.2	5.2
IBO/LBO	.6	.9	1.9
LBO	12.4	31.6	33.5
LBO/MAVO	4.5	7.3	9.8
MAVO	24.7	26.9	25.1
MAVO/HAVO	10.4	6.7	10.0
HAVO	21.2	11.3	5.6
HAVO/VWO	15.9	8.6	4.5
VWO	7.7	2.2	1.2
OVERIG	2.3	2.4	3.1
N	2259	1831	2266

De verschillen in adviezen tussen leerlingen uit de OVB-doelgroepen en de overige leerlingen zijn overduidelijk. De adviezen aan de 1.25- en de 1.90-leerlingen zijn veel lager dan de adviezen aan de 1.00-leerlingen. Van de 1.00-leerlingen krijgt 45 procent een HAVO- of VWO-advies, bij de 1.25-leerlingen is dat percentage 22 en bij de 1.90-leerlingen slechts 11. Eén op de drie leerlingen uit de OVB-doelgroepen krijgt een LBO-advies, tegen één op de acht van de 1.00-leerlingen. De percentages MAVO-adviezen zijn bij de drie wegingsfactoren ongeveer gelijk. Opvallend is het hoge percentage 1.90-leerlingen dat een IBO- of IBO/LBO-advies krijgt. Eén op de 14 van deze leerlingen heeft volgens de basisschool in het voortgezet onderwijs (wellicht) extra aan-

Tabel 3
Gemiddelde T-scores voor taal en rekenen per advies, uitgesplitst naar wegingsfactor

	TAAL			REKENEN		
	1.00	1.25	1.90	1.00	1.25	1.90
IBO	*	36.8	32.1	*	36.1	32.8
IBO/LBO	39.0	39.4	32.3	37.9	35.0	36.3
LBO	44.8	43.5	35.8	44.0	43.4	38.1
LBO/MAVO	44.8	46.2	39.0	45.1	45.9	40.7
MAVO	50.1	49.2	40.9	48.7	49.2	44.2
MAVO/HAVO	51.9	53.8	44.0	52.0	52.2	47.9
HAVO	55.6	55.7	47.8	56.6	56.3	52.4
HAVO/VWO	58.0	56.9	48.4	58.3	57.0	53.8
VWO	61.3	59.8	53.2	60.9	62.7	56.9

* < 10

dacht nodig in de vorm van individueel onderwijs.

Aan 1.90-leerlingen wordt vaker een laag advies gegeven dan aan 1.00- en 1.25-leerlingen. Toch lijken de verschillen in adviezen tussen de groepen minder opzienbarend dan de verschillen in taal- en rekenprestaties. Dat zou erop kunnen wijzen dat de relatie tussen prestaties en advies per wegingsfactor verschilt. Of dat zo is, hebben we onderzocht door voor de drie wegingsfactoren de gemiddelde taal- en rekenscores per advies te berekenen. De resultaten daarvan staan in Tabel 3.

De gemiddelde taal- en rekenprestaties per advies vertonen een opvallend regelmatig verloop. Hoe hoger het advies, hoe hoger de gemiddelde taal- en rekenscores. Zelfs de gemiddelde scores van de gecombineerde adviezen liggen vrij precies tussen het hogere en lagere advies in.

De gemiddelde scores van de 1.00-leerlingen en van de 1.25-leerlingen ontlopen elkaar bij de meeste adviezen niet zoveel. Bij sommige adviezen zijn de scores van de 1.00-leerlingen wat hoger, bij andere adviezen die van de 1.25-leerlingen. De gemiddelde scores van de allochtone doelgroepleerlingen wijken daarentegen zeer sterk af van die van de 1.00- en 1.25-leerlingen. Bij eenzelfde advies hebben de 1.90-leerlingen consequent veel lagere scores dan de leerlingen uit de twee andere groepen.

Betekent dit nu ook dat bij gelijke prestaties 1.90-leerlingen een hoger advies krijgen dan 1.00- en 1.25-leerlingen? En zo ja, welke gevolgen hebben de relatief hoge adviezen voor de verdere schoolloopbaan in het voortgezet onderwijs? Over die vragen gaat het tweede deel van deze bijdrage.

2.2 Hogere adviezen aan allochtone leerlingen?

Op het eerste gezicht lijken de gemiddeld lagere prestaties per advies bij de 1.90-leerlingen erop te wijzen dat er bij de advisering aan 1.90-leerlingen lagere eisen worden gesteld dan aan 1.00- en 1.25-leerlingen. Dat zou betekenen dat 1.90-leerlingen dus relatief hoge adviezen krijgen. Toch mag die conclusie op basis van deze voorstelling van zaken niet worden getrokken. Door de gemiddelde prestaties te berekenen gegeven een bepaald advies - de prestaties te corrigeren voor het advies - wordt namelijk impliciet een model gehanteerd waarin het advies van invloed is op de prestaties (zie Mulder & Pijl, 1992). Een dergelijk model is natuurlijk irreëel aangezien het advies de prestaties niet beïnvloedt, maar de prestaties het advies. De vraag of 1.90-leerlingen relatief hoge adviezen krijgen, kan alleen worden beantwoord door de adviezen aan 1.00-, 1.25- en 1.90-leerlingen met dezelfde taal- en rekenprestaties met elkaar te vergelijken. We hebben dat gedaan met behulp van covariantie-analyse. Bij deze analyse wordt het effect van een variabele op een andere variabele onderzocht, nadat de scores op de afhankelijke variabele zijn 'gezuiverd' van de invloed van andere, op dat moment niet relevant geachte, variabelen. In dit verband gaat het om het effect van de wegingsfactor op het advies, nadat is gecontroleerd voor de taal- en rekenscores. Het effect wordt weergegeven als de gemiddelde afwijking van de betreffende wegingsfactor ten opzichte van het gemiddelde advies van de totale groep leerlingen. Wijkt de

score van een bepaalde wegingsfactor af van het gemiddelde, dan heeft die wegingsfactor, ook na controle voor de prestaties, nog een zelfstandig effect op het advies. Als het zo is dat 1.90-leerlingen bij gelijke prestaties hogere adviezen krijgen dan de 1.00- en 1.25-leerlingen, dan zal de afwijkingsscore van wegingsfactor 1.90 ten opzichte van het totale gemiddelde advies groter zijn dan de afwijkingsscore van de twee andere wegingsfactoren.

In de analyses zijn de adviezen geschaald volgens de leerjarenladder van Bosker en Van der Velden (1989). Het uitgangspunt bij de leerjarenladder is dat 6-VWO het hoogst bereikbare niveau in het voortgezet onderwijs is. Aan de voltooiing van VWO-6 wordt de waarde 12 toegekend. De positie van een leerling wordt weergegeven als de mate waarin die top van het onderwijssysteem is bereikt in termen van leerjaren. Het minimum van de leerjarenladder wordt op één gesteld en houdt in dat er voortgezet speciaal onderwijs is geadviseerd. Een leerling die een VWO-advies krijgt, 'springt' meteen naar de zesde tree. De schaling van de adviezen volgens de leerjarenladder ziet er als volgt uit:

SO	1	MAVO	4
IBO	2	MAVO/HAVO	4,5
IBO/LBO	2,5	HAVO	5
LBO	3	HAVO/VWO	5,5
LBO/MAVO	3,5	VWO	6

De resultatèn van de covariantie-analyse staan in Tabel 4. Voorafgaand aan de analyse is getoetst op mogelijke interactie-effecten, maar deze bleken niet significant. De gemiddelde adviesscore van de totale groep (de Grand Mean) bedraagt 4.14. De getallen in Tabel 4 geven per wegingsfactor de afwijking ten opzichte van dat gemiddelde aan. In de linker kolom staat de afwijking wanneer niet wordt gecontroleerd voor de taal- en rekenprestaties, in de rechter kolom staat de afwijking wanneer daar wel voor wordt gecontroleerd.

Wanneer geen rekening wordt gehouden met de taal- en rekenprestaties van de leerlingen krijgen de 1.00-leerlingen de hoogste adviezen en de 1.90-leerlingen de laagste. De 1.25-leerlingen zitten daar tussenin. Na controle voor de prestaties veranderen de afwijkingsscores. Bij gelijke prestaties zijn er geen verschillen meer tussen de adviezen aan 1.00- en 1.90-leerlingen. De adviezen aan de 1.25-leerlingen zijn duidelijk lager. Het feit dat de bèta veel kleiner is dan de èta geeft aan dat de samenhang tussen wegingsfactor en advies voor een groot deel door de prestaties wordt bepaald. Toch blijft er ook na controle voor de prestaties een rechtstreeks effect van wegingsfactor op advies over. Op basis van de gegevens in de tabel kan de vraag of 1.90-leerlingen relatief hoge adviezen krijgen worden beantwoord met: ja, maar alleen in vergelijking met leerlingen die ook tot de OVB-doelgroepen behoren, dus alleen vergeleken met leerlingen met een vergelijkbare sociaal-economische achtergrond. In vergelijking met de 1.00-leerlingen is er geen enkel verschil.

Tabel 4

Afwijkingsscores per OVB-wegingsfactor van het gemiddelde advies, zonder en met controle voor taal- en rekenprestaties. Grand Mean= 4.14

wegingsfactor	zonder controle	met controle
1.00	.30	.12
1.25	–.27	–.17
1.90	–.47	.11
èta	.30	
bèta		.14

2.3 De gevolgen van het advies voor de loopbaan in het voortgezet onderwijs

Uit schoolloopbaanonderzoek is bekend dat de doorstroom naar het type voortgezet onderwijs nauw samenhangt met het sociaal milieu waaruit de leerling komt. Er vindt een scheiding naar milieu plaats waarbij de OVB-doelgroepleerlingen overwegend naar LBO en MAVO gaan en de 1.00-leerlingen veel vaker op HAVO of VWO terechtkomen. Door de relatief hoge adviezen is het gemiddelde prestatieniveau van de 1.90-leerlingen aan het begin van het voortgezet onderwijs lager dan dat van de 1.25-leerlingen op hetzelfde schooltype. Het is niet ondenkbaar dat 1.90-leerlingen daardoor vaker blijven zitten of vaker naar een ander (lager) schooltype gaan. Het effect van het relatief hoge advies is echter niet zonder meer vast te stellen. Een complicerende factor is namelijk dat allochtone leerlingen vaker dan autochtone leerlingen voor een schooltype kiezen dat hoger is dan het advies (Mulder, 1991; Mulder & Tesser, 1992). En ook een hogere keuze heeft tot gevolg dat het prestatieniveau

van 1.90-leerlingen lager is dan dat van hun Nederlandse klasgenoten, met een mogelijk eigen effect op de verdere schoolloopbaan.

Om toch enig zicht te krijgen op het gevolg van de relatief hoge adviezen aan 1.90-leerlingen is de volgende weg bewandeld. Er is uitgegaan van het gegeven dat in het eerste leerjaar van het voortgezet onderwijs de gemiddelde taal- en rekenprestaties van de 1.90-leerlingen binnen elk schooltype lager zijn dan die van de 1.00- en 1.25-leerlingen. De lagere prestaties kunnen zowel een gevolg zijn van de relatief hoge adviezen, als van de keuze voor een hoger schooltype dan is geadviseerd. Eerst is onderzocht of die lagere prestaties aan het begin van het voortgezet onderwijs in het tweede leerjaar een lagere onderwijspositie tot gevolg hebben. Hierbij is de onderwijspositie weer geoperationaliseerd in termen van de leerjarenladder. Wanneer de 1.90-leerlingen in het tweede leerjaar een lagere onderwijspositie hebben dan de 1.00- of 1.25-leerlingen *die in het eerste leerjaar in een vergelijkbaar schooltype zaten*, dan zou dat kunnen komen door de relatief hoge adviezen en/of door het feit dat allochtonen vaker voor een hoger schooltype kiezen dan is geadviseerd. Wanneer er tussen de drie wegingsfactoren echter géén verschillen in gemiddelde onderwijspositie zijn, dan luidt de conclusie dat de relatief hoge adviezen (nog) geen zichtbare nadelige gevolgen hebben.

De vraag of er in het tweede leerjaar verschillen in onderwijspositie tussen de wegingsfactoren bestaan, is onderzocht met behulp van regressie-analyse. De wegingsfactor is als dummy-variabele in de analyse opgenomen. Daarbij is van elke wegingsfactor een afzonderlijke variabele gemaakt die twee waarden kan aannemen: een 1 als de leerling de betreffende wegingsfactor heeft en een 0 als dit niet het geval is. In de regressie-analyses met dummies doet één wegingsfactor niet mee, namelijk de 1.00-leerlingen. Deze groep fungeert als vergelijkingsgroep waartegen de 1.25- en 1.90-leerlingen worden afgezet. De regressiecoëfficiënten van de dummies representeren de verschillen in gemiddelde onderwijspositie van de betreffende wegingsfactor ten opzichte van de gemiddelde onderwijspositie van de vergelijkingsgroep. De analyses zijn per schooltype uitgevoerd, telkens op de leerlingen die in het eerste leerjaar op eenzelfde schooltype zaten. Om de significatiewaarden van de effecten van de wegingsfactoren onderling te kunnen vergelijken, zijn de aantallen 1.00-, 1.25- en 1.90-leerlingen *binnen* eenzelfde schooltype gelijk gesteld. De resultaten van de analyses staan in Tabel 5.

Tabel 5

Effecten van wegingsfactor op de onderwijspositie in het tweede leerjaar, vergelijkingsgroep: 1.00-leerlingen (ongestandaardiseerde regressie-coëfficiënten)

	1.25	1.90
IBO	n.s.	.21
IBO/LBO	n.s.	–.45
LBO	n.s.	n.s.
LBO/MAVO	n.s.	n.s.
MAVO	n.s.	–.11
MAVO/HAVO	n.s.	–.23
HAVO	–.52	*
HAVO/VWO	–.22	–.34
VWO	n.s.	*
MAVO/HAVO/VWO	–.30	–.36

* < 10

Voor de 1.25-leerlingen die in het eerste leerjaar in een VWO-klas zaten of in een klas lager dan HAVO, geldt dat hun onderwijspositie in het tweede leerjaar niet afwijkt van die van de 1.00-leerlingen. De 1.25-leerlingen die in het eerste leerjaar in een HAVO-, HAVO/VWO of MAVO/HAVO/VWO-klas zaten, hebben in het tweede leerjaar gemiddeld een lagere positie dan de 1.00-leerlingen. Op de hogere schooltypen (HAVO-VWO-niveau) doen de 1.00-leerlingen het dus beter dan de 1.25-leerlingen.

De onderwijspositie van 1.90-leerlingen in een LBO- of LBO/MAVO-klas wijkt niet significant van die van de 1.00-leerlingen af. Dat is wel het geval bij de overige schooltypen. De 1.90-leerlingen in IBO-klassen hebben in het tweede leerjaar een hogere onderwijspositie bereikt dan de 1.00-leerlingen; 1.90-leerlingen die in een IBO/LBO-, MAVO-, MAVO/HAVO- HAVO/VWO- of MAVO/HAVO/VWO-klas zaten, hebben gemiddeld een lagere onderwijspositie bereikt.

Over de onderwijspositie van de 1.90-leerlingen in een HAVO- of VWO-klas kunnen we geen uitspraken doen. Het totaal aantal leerlingen in die klassen is daarvoor te klein.

In het tweede leerjaar is de onderwijspositie

van 1.90-leerlingen dus ongunstiger dan die van 1.00- en 1.25-leerlingen die in het eerste jaar op hetzelfde schooltype zaten. De 1.90-leerlingen zijn vaker blijven zitten, vaker naar een lager schooltype overgestapt of minder vaak naar een hoger schooltype gegaan. Dat zou te maken kunnen hebben met de relatief hoge adviezen of met het feit dat ze vaker voor een hoger schooltype kiezen dan is geadviseerd. Voor het effect van de schoolkeuze is gecontroleerd door leerlingen met elkaar te vergelijken die op dezelfde manier van het advies zijn afgeweken. Daartoe zijn drie groepen leerlingen onderscheiden: leerlingen die het advies hebben opgevolgd, leerlingen die lager hebben gekozen dan het advies en leerlingen die voor een hoger schooltype hebben gekozen dan geadviseerd. Het opvolgen van het advies is ook als dummy-variabele in de vergelijking opgenomen. Leerlingen die het advies hebben opgevolgd fungeren als vergelijkingsgroep, leerlingen die hebben gekozen voor een hoger dan wel lager schooltype worden tegen de vergelijkingsgroep afgezet. Verdwijnt het effect van de wegingsfactor na toevoeging van de variabele 'afwijking van het advies' dan kan worden geconcludeerd dat de lagere onderwijspositie bij het betreffende schooltype is veroorzaakt door het hoger of lager kiezen dan het advies. In Tabel 6 is te zien bij welke schooltypen het afwijken van het advies een rol speelt bij de onderwijspositie van de leerlingen. Bij MAVO/HAVO/VWO-klassen kan niet voor de keuze worden gecontroleerd, omdat het advies vrijwel altijd onderdeel uitmaakt van het gekozen schooltype.

Bij de 1.90-leerlingen verdwijnt bij vrijwel alle schooltypen het verschil in onderwijspositie ten opzichte van de 1.00-leerlingen na controle voor het afwijken van het advies. Wanneer we leerlingen met elkaar vergelijken die op dezelfde manier van het advies zijn afgeweken is er geen verschil meer in onderwijspositie. Dat geldt voor 1.90-leerlingen die in het eerste jaar in een IBO/LBO-, MAVO- of MAVO/HAVO-klas zaten. De lagere onderwijspositie van deze leerlingen is dus te wijten aan hun schoolkeuze. Hun *te hoge* schoolkeuze, want dat blijkt op alle schooltypen een negatief effect te hebben op de onderwijspositie in het tweede leerjaar. Bij IBO en HAVO/VWO blijft ook na controle voor het afwijken van het advies een effect van de wegingsfactor bestaan. Bij het IBO is het onmogelijk om hoger te kiezen dan het advies, en een lagere keuze heeft geen enkel effect. Bij HAVO/VWO heeft de schoolkeuze wel invloed (het effect van de wegingsfactor wordt immers kleiner), maar toch is de hogere keuze niet de hele verklaring voor de lagere onderwijspositie van de 1.90-leerlingen.

Bij de 1.25-leerlingen die in het eerste jaar in een HAVO-klas zaten, verdwijnt het verschil ten opzichte van de 1.00-leerlingen na controle voor de keuze. Ook hier is de hoge keuze de verklaring voor de lagere onderwijspositie in het tweede jaar. Bij HAVO/VWO geldt dat niet. Net als bij de 1.90-leerlingen blijft er bij dit schooltype nog een effect van de wegingsfactor bestaan.

Bij HAVO/VWO-leerlingen is er blijkbaar nog een andere factor in het geding. Of die andere

Tabel 6

Effecten van wegingsfactor op de onderwijspositie in het tweede leerjaar, zonder en met controle voor het afwijken van het advies, vergelijkingsgroep: 1.00-leerlingen (ongestandaardiseerde regressie-coëfficiënten)

	zonder controle		met controle		hogere	lagere
	1.25	1.90	1.25	1.90	keuze	keuze
IBO	n.s.	.21	n.s.	.21	—	n.s.
IBO/LBO	n.s.	–.45	n.s.	n.s.	–.84	n.s.
LBO	n.s.	n.s.	n.s.	n.s.	–.36	n.s.
LBO/MAVO	n.s.	n.s.	n.s.	n.s.	–.18	n.s.
MAVO	n.s.	–.11	n.s.	n.s.	–.16	.32
MAVO/HAVO	n.s.	–.23	n.s.	n.s.	–.29	.48
HAVO	–.52	*	n.s.	*	–.55	n.s.
HAVO/VWO	–.22	–.34	–.19	–.25	–.35	n.s.
VWO	n.s.	*	n.s.	*	–.51	—
MAVO/HAVO/VWO	–.30	–.36				

* < 10

Tabel 7

Effecten van wegingsfactor op de onderwijspositie in het tweede leerjaar, na controle voor het afwijken van het advies en de taal- en rekenprestaties, vergelijkingsgroep: 1.00-leerlingen (ongestandaardiseerde regressie-coëfficiënten)

	1.25	1.90	hogere keuze	lagere keuze	taal	rekenen
IBO	n.s.	.22	—	n.s.	n.s.	n.s.
IBO/LBO	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
LBO	n.s.	n.s.	−.31	n.s.	.005	.005
LBO/MAVO	n.s.	n.s.	−.18	n.s.	n.s.	n.s.
MAVO	n.s.	n.s.	−.12	.29	.006	n.s.
MAVO/HAVO	n.s.	n.s.	−.25	n.s.	n.s.	.02
HAVO	n.s.	*	−.44	n.s.	.06	n.s.
HAVO/VWO	−.17	n.s.	−.26	n.s.	.01	.01
VWO	n.s.	*	−.50	—	n.s.	n.s.
MAVO/HAVO/VWO	−.27	n.s.			.02	.03

* < 10

factor het advies betreft kunnen we nagaan door naast de wegingsfactor en het afwijken van het advies, ook de taal- en rekenscores van de leerlingen in de analyses op te nemen. Wanneer dat tot gevolg heeft dat het effect van de wegingsfactor verdwijnt, dan wijst dat erop dat behalve de afwijking van het advies ook het advies zelf verantwoordelijk is voor de lagere onderwijspositie. Het effect van de prestaties kan immers niet meer worden toegeschreven aan de hogere schoolkeuze, omdat we daarvoor controleren. Het moet te maken hebben met het feit dat ook leerlingen die op dezelfde manier van het advies zijn afgeweken nog met een verschillend prestatieniveau het VO binnenkomen. En dus relatief hoge of lage adviezen hebben gehad.

De resultaten van de analyses met de taal- en rekenprestaties en de schoolkeuze staan in Tabel 7.

Bij de 1.90-leerlingen die in het eerste leerjaar in een HAVO/VWO-klas zaten, verdwijnt inderdaad het effect van de wegingsfactor helemaal na controle voor de taal- en rekenprestaties. Bij hen speelt naast de afwijking van het advies ook het advies zelf een rol bij hun lagere onderwijspositie in het tweede leerjaar. Zij hebben blijkbaar gekozen voor een hoger schooltype dan was geadviseerd, terwijl dat advies al relatief hoog was.

Bij de 1.25-leerlingen hebben de taal- en rekenprestaties geen effect op het verschil met de 1.00-leerlingen. Dat was ook niet te verwachten want binnen de schooltypen waren de prestatieverschillen tussen de 1.00- en 1.25-leerlingen minimaal. Waarom zij in het tweede leerjaar gemiddeld een lagere onderwijspositie hebben dan de 1.00-leerlingen is niet geheel duidelijk.

3 Conclusie

Bij gelijke prestaties zijn de adviezen aan 1.90-leerlingen hoger dan die aan 1.25-leerlingen, maar tussen de adviezen aan 1.90- en 1.00-leerlingen bestaat geen verschil. Dat betekent dat de bewering dat allochtone leerlingen relatief hoge adviezen krijgen in zijn algemeenheid niet houdbaar is. De bewering geldt slechts voorzover het allochtone en autochtone leerlingen met vergelijkbare sociale achtergrond betreft. Voor de praktische consequenties maakt die beperking echter niet erg veel uit. In het voortgezet onderwijs komen allochtone en Nederlandse leerlingen van vergelijkbaar milieu op hetzelfde schooltype bij elkaar in de klas. Binnen die schooltypen zouden de hogere adviezen aan allochtonen een negatief effect kunnen hebben op de verdere schoolloopbaan. We hebben, met uitzondering van het HAVO/VWO-advies, geen nadelige gevolgen vast kunnen stellen van de relatief hoge adviezen aan allochtone leerlingen voor hun onderwijspositie in het tweede leerjaar. Het kan zijn dat de verschillen tussen de adviezen aan autochtonen en aan allochtonen met gelijke prestaties te klein zijn om zo'n effect teweeg te brengen. Het kan echter ook betekenen dat de adviserende leerkracht of directeur een juiste inschatting heeft gemaakt van het niveau van het kind, en dat hij of zij meer in het kind ziet dan uit

toetsprestaties kan worden afgelezen. Bij de advisering spelen immers veel andere factoren een rol, zoals motivatie en doorzettingsvermogen van het kind en de te verwachten ondersteuning van de ouders.

Dat de onderwijspositie van de 1.90-leerlingen in het tweede leerjaar ongunstiger is dan die van 1.00- en 1.25-leerlingen die in het eerste jaar op hetzelfde schooltype zaten, is voornamelijk een gevolg van het feit dat 1.90-leerlingen relatief vaak van het advies afwijken en voor een hoger schooltype kiezen dan is geadviseerd. Het niveau dat veel allochtone ouders voor hun kind kiezen is blijkbaar te hoog gegrepen. Waarom veel allochtone ouders het advies niet opvolgen is nog steeds niet helemaal duidelijk. Een van de redenen zou kunnen zijn dat zij een onjuiste inschatting maken van de mogelijkheden van hun kind. Wanneer ouders zelf de Nederlandse taal niet of nauwelijks beheersen en hun kind veel beter Nederlands spreekt dan zijzelf kan dat gemakkelijk leiden tot een overschatting van het niveau van hun kind. Dat geldt dan met name voor Turkse en Marokkaanse ouders.

In ander onderzoek wordt het hoge aspiratieniveau van allochtone ouders genoemd. Veel van hen zijn naar Nederland gekomen om een beter bestaan op te bouwen. Voor henzelf is dat vaak niet of maar ten dele gelukt. Het is mogelijk dat ze voor hun kind een betere toekomst wensen en dat een goede opleiding daartoe de eerste stap is. Ten slotte zal ook de onbekendheid met het Nederlandse onderwijssysteem een belangrijke factor zijn. Dat er grote behoefte is aan informatie over de schoolkeuze blijkt uit het feit dat allochtone ouders in groten getale afkomen op informatie-bijeenkomsten die in de eigen taal worden gegeven. Deze bijeenkomsten, die nu nog op slechts enkele plaatsen in het land worden georganiseerd, verdienen zeker navolging, want met voorlichting in het Nederlands wordt maar een klein deel van de allochtone ouders bereikt.

Literatuur

Bosker, R.J., & Velden, R.K.W. van der (1989). Schooleffecten en rendementen. In H.P.M. Creemers e.a. (red.). *Bijdragen aan de onderwijsresearch*, nr. 25. Lisse: Swets & Zeitlinger.

Driessen, G.W.J.M. (1991). Discrepanties tussen toetsresultaten en doorstroomniveau. Positieve discriminatie bij de overgang basisonderwijs – voortgezet onderwijs? *Pedagogische Studiën, 68*, 27-35.

Jong, M.-J. de (1987). *Herkomst, kennis en kansen. Allochtone en autochtone leerlingen tijdens de overgang van basis- naar voortgezet onderwijs.* Lisse: Swets & Zeitlinger.

Mulder, L. (1991). *De overgang van basis- naar voortgezet onderwijs van de OVB-doelgroepen. Advies, schoolkeuze en rapportcijfers in het eerste leerjaar.* Nijmegen/Groningen: ITS/RION.

Mulder, L., & Tesser, P. (1992). *De schoolkeuzen van allochtone leerlingen.* Nijmegen: ITS.

Mulder, L., & Pijl, Y. (1992). *De onderwijspositie van leerlingen uit de OVB-doelgroepen na twee jaar voortgezet onderwijs.* Nijmegen/Groningen: ITS/RION.

Tesser, P., Mulder, L., & Werf, M. van der (1991). *De eerste fase van de longitudinale OVB-cohortonderzoeken.* Het leerlingenonderzoek. Nijmegen/Groningen: ITS/RION.

Manuscript aanvaard 11-1-1993

Auteur

L. Mulder is werkzaam op het Instituut voor Toegepaste Sociale Wetenschappen te Nijmegen en is sinds 1987 betrokken bij de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid.

Adres: ITS, Postbus 9048, 6500 KJ Nijmegen

Abstract

The effect of the recommendation on the secondary school careers of non-indigenous pupils belonging to the OVB-target groups

L. Mulder. Pedagogische Studiën, 1993, 70, 242-251.

Secondary school career recommendations are for the greater part based on schoolachievements. However, the relationship between achievement and recommendation is not entirely unambiguous. In recent years various researchers have come to the conclusion that the recommendations made to non-

indigenous pupils are relatively high. The question arises what the consequence of this is for the further secondary school career of non-indigenous pupils. This article focuses on this question. Data from the National Evaluation on Dutch Educational Priority Policy (OVB) formed the basis for the analyses. The results show that non-indigenous pupils indeed get relatively high recommendations, but only in comparison with indigenous pupils of comparable social background. The relatively high recommendations don't seem to be harmful to the further school career, at least not in the first two years of secondary school. After two years of secondary school the careers of non-indigenous pupils are worse than those of other pupils, but this is mainly due to the fact that non-indigenous pupils choose more often a higher school type than has been recommended.

Binnen en buiten Onderwijsvoorrangsgebieden. Sociaal-etnische achtergronden en onderwijspositie van leerlingen in het voortgezet onderwijs*

G. Driessen**

Samenvatting

Het gebiedenbeleid vormt een van de componenten van het Onderwijsvoorrangsbeleid. Binnen onderwijsvoorrangsgebieden (OVG's) is sprake van een cumulatie van achterstanden. Om deze te verminderen gaan basisscholen, scholen voor voortgezet onderwijs en welzijnsinstellingen samenwerkingsrelaties aan. Over de situatie van leerlingen op scholen voor voortgezet onderwijs binnen de OVG's is nog weinig bekend. De centrale vraag van deze bijdrage luidt dan ook: waarin onderscheiden v.o.-leerlingen binnen OVG's zich van v.o.-leerlingen buiten OVG's? Deze vraag wordt beantwoord met behulp van gegevens die zijn verzameld in het CBS-cohortonderzoek VOCL'89. Het blijkt dat er zich verschillen voordoen tussen OVG's en niet-OVG's met betrekking tot de sociaal-etnische achtergronden van de leerlingenpopulaties, het niveau van de schooltypen en de toetsprestaties. Echter, de verschillen in prestaties kunnen nagenoeg geheel worden herleid tot verschillen in schooltype en sociaal-etnische achtergronden. Na controle voor de invloed van de laatste twee factoren is het verschil in prestaties tussen leerlingen binnen en buiten gebieden nog slechts uiterst gering.

Inleiding

Teneinde onderwijsachterstanden die het gevolg zijn van sociale, economische en culturele omstandigheden op te heffen dan wel te verminderen, wordt het Onderwijsvoorrangsbeleid (OVB) uitgevoerd. Binnen dit beleid kunnen twee componenten worden onderscheiden, namelijk het formatiebeleid en het gebiedenbeleid. Het formatiebeleid houdt in dat in het basisonderwijs de personeelsformatie van de scholen afhankelijk wordt gesteld van de sociaal-etnische samenstelling van hun leerlingenpopulatie. In het voortgezet onderwijs komen scholen in aanmerking voor extra formatie in verband met de aanwezigheid van allochtone leerlingen. De tweede component van het OVB is het gebiedenbeleid. Binnen onderwijsvoorrangsgebieden - dit zijn geografisch bepaalde eenheden waarin sprake is van een cumulatie van achterstanden - worden samenwerkingsverbanden aangegaan tussen basisscholen, scholen voor voortgezet onderwijs en welzijnsinstellingen (zoals peuterspeelzalen, bibliotheken). In 1986 zijn 70 van dergelijke OVG's van start gegaan. Zij hebben tot doel concrete vormen van onderwijsachterstandsbestrijding te ontwikkelen en in praktijk te brengen. Om dit doel te verwezenlijken krijgen de gebieden subsidie van de ministeries van Onderwijs en Wetenschappen en Welzijn, Volksgezondheid en Cultuur.

Tot op heden is er in beperkte mate onderzoek gedaan naar het functioneren van de gebieden. Verwezen wordt op deze plaats naar onder andere Research voor Beleid (1987a, 1987b), Tesser, Van der Werf, Mulder en Weide (1989), Vermeulen en Van der Werf (1989), Verhaak en Brandsma (1991) en Tesser, Mulder en Van der Werf (1991). Genoemde studies hebben onder meer betrekking op de organisatie van OVG's, op de samenwer-

* De gegevens die gebruikt zijn voor de hier beschreven analyses zijn afkomstig uit het cohortonderzoek VOCL'89. De data zijn verzameld en beschikbaar gesteld door het Centraal Bureau voor de Statistiek (CBS). De uitvoering van de oorspronkelijke analyses is mogelijk gemaakt door een subsidie van het Instituut voor Onderzoek van het Onderwijs (SVO) te 's-Gravenhage.

** Met dank aan L. Mulder, G. W. Meijnen en J. Kloprogge voor hun commentaar op een eerdere versie van dit artikel en H. Migchielsen voor het verstrekken van enige achtergrondinformatie.

king tussen onderwijsinstanties en welzijnsinstellingen, op de inhoud van de gemaakte gebiedsplannen en op de prestaties van leerlingen in het basisonderwijs binnen en buiten gebieden. Waar nog nauwelijks zicht op is, is de situatie van de scholen voor voortgezet onderwijs binnen de gebieden. Meer specifiek: onbekend is in welke opzichten deze scholen binnen gebieden nu precies verschillen van scholen buiten gebieden, bijvoorbeeld voor wat betreft de samenstelling van hun leerlingenpopulaties en de door die leerlingen geleverde prestaties.

Recentelijk zijn er in het kader van het CBS-cohortonderzoek VOCL'89 gegevens beschikbaar gekomen die voor een deel in deze lacune kunnen voorzien. Heel concreet gaat het in dat onderzoek om gegevens over sociaal-etnische achtergrondkenmerken van de leerlingen, over hun bereikte onderwijsniveau en over de daarbinnen behaalde prestaties. Daarnaast is bekend of de scholen waarop de leerlingen zitten participeren aan een onderwijsvoorrangsgebied. Voor de duidelijkheid: bij VOCL'89 gaat het om het voortgezet onderwijs, dit in tegenstelling tot eerder onderzoek op leerlingniveau dat betrekking had op de situatie in het basisonderwijs.

Een van de aspecten van de OVB-evaluatie behelst de vraag naar effecten van participatie aan een OVG. Omdat het bij het onderzoek waarover in dit artikel wordt gerapporteerd pas om een eerste meting gaat, valt daar de facto nog niets over te zeggen. Bovendien behoren voor een goede interpretatie te zijner tijd ook gegevens over ondernomen OVG-activiteiten in de analyses te worden betrokken. Voor het vervolg van de OVB-evaluatie is het echter zinvol nu al inzicht te krijgen in de kenmerken van de leerlingenpopulaties binnen en buiten gebieden. In het licht van het bovenstaande dient 'OVG-participatie' vooralsnog niet te worden opgevat in de zin van dat er extra activiteiten voor OVB-doelgroepen worden ondernomen, dan wel dat het een indicatie geeft van het feit dat er op de scholen die participeren, sprake is van een extra verzwaring van de achterstandsproblematiek als gevolg van een cumulatie van problemen.

In de hierna volgende analyses staan drie vragen centraal:

1. Bestaan er verschillen in sociaal-etnische achtergrondkenmerken tussen leerlingen op scholen voor voortgezet onderwijs die al dan niet deelnemen aan onderwijsvoorrangsgebieden?
2. Bestaan er verschillen in onderwijspositie en toetsprestaties tussen leerlingen op scholen voor voortgezet onderwijs die al dan niet deelnemen aan onderwijsvoorrangsgebieden?
3. Bestaan er verschillen in de verklaring van verschillen in onderwijspositie en toetsprestaties tussen leerlingen op scholen voor voortgezet onderwijs die al dan niet deelnemen aan onderwijsvoorrangsgebieden?

De analyses worden uitgevoerd op leerlingniveau; ze zijn van beschrijvende en exploratieve aard.

De geformuleerde vragen hebben misschien op het eerste oog het karakter van 'open deuren'; om duidelijk te maken dat dit niet het geval is, is het belangrijk kort in te gaan op de ontstaansgeschiedenis van de OVG's rond 1985.

Binnen een OVG dienen ten minste 10 basisscholen, drie categorieën van voortgezet onderwijs en één welzijnsinstelling samen te werken. Aan de basisscholen werd de eis gesteld dat zij een bepaalde mate van achterstand moesten kennen, uitgedrukt in de zogenaamde SE-factor (sociaal-economische factor). Voor het voortgezet onderwijs gold dat er ten minste twee soorten van lager beroepsonderwijs en één soort algemeen vormend onderwijs, c.q. voorbereidend wetenschappelijk onderwijs vertegenwoordigd moesten zijn. Daarnaast diende er een bepaalde doorstroomrelatie te bestaan tussen de participerende basisscholen en scholen voor voortgezet onderwijs: ten minste 20 procent van de leerlingen die de deelnemende basisscholen verlieten moesten naar de deelnemende brugklassen doorstromen en het leerlingenbestand in de brugklassen diende voor ten minste 5 procent afkomstig te zijn van de deelnemende basisscholen (Ministerie van Onderwijs en Wetenschappen, 1985b).

Duidelijk zal zijn dat deze criteria een aanzienlijke speelruimte over laten met betrekking tot de sociaal-economische en etnische samenstelling van de v.o.-scholen. Aan de 20, respectievelijk 5 procent doorstromers werden namelijk geen eisen gesteld; dat konden net zo goed

1.90- als 1.00-leerlingen zijn. En dat laatste gold evenzeer voor de resterende 80, respectievelijk 95 procent. Verder is het van belang dat er zich oorspronkelijk 116 potentiële gebieden hebben aangemeld, waarvan er uiteindelijk (slechts) 70 zijn erkend. Als criterium daarbij werd de zwaarte van de achterstand van het gebied gehanteerd. Bij de bepaling van de achterstand is uitgegaan van het aantal leerlingen dat op 16 september 1985 bij de scholen stond ingeschreven (Ministerie van Onderwijs en Wetenschappen, 1985a; Stimulans, 1985).

Samenvattend is er dus sprake van behoorlijke marges met betrekking tot de samenstelling van de schoolpopulaties in het voortgezet onderwijs (en dat geldt met name in het geval van brede scholengemeenschappen). Voorts zijn er destijds misschien scholen niet in een OVG opgenomen omdat het betreffende gebied net iets minder achterstand vertoonde dan een ander gebied en bovendien zijn er sedert de vorming van de gebieden alweer een aantal jaren verstreken. Al met al dus voldoende reden om na te gaan hoe de situatie in het voortgezet onderwijs er momenteel uit ziet.

1 Steekproef en variabelen

De gebruikte gegevens zijn oorspronkelijk verzameld ten behoeve van het cohortonderzoek VOCL'89. Het gaat hierbij om een onderzoek waarbij leerlingen gedurende hun loopbaan in het voortgezet onderwijs worden gevolgd. Centraal daarbij staat de relatie tussen onderwijspositie en sociaal-etnische achtergronden. De eerste gegevensverzameling heeft plaatsgevonden in het schooljaar 1989/1990. In dat jaar zijn onder andere toetsen afgenomen bij leerlingen in het eerste leerjaar voortgezet onderwijs en zijn achtergrondkenmerken verzameld bij hun ouders.

De populatie die het uitgangspunt vormde voor het onderzoek bestond uit 1818 scholen voor voortgezet onderwijs uit de categorieën ibo, lbo, mavo, havo, vwo, smalle scholengemeenschap en brede scholengemeenschap. Op die scholen zaten in totaal 185.198 leerlingen in het eerste leerjaar. Bij de steekproeftrekking is gekozen voor een tweetrapssteekproef, hetgeen inhoudt dat eerst scholen zijn getrokken en vervolgens daarbinnen klassen. De steekproef bestaat uit 19.524 leerlingen van 381 scholen. Van de deelnemende scholen zijn er 73 afkomstig uit een OVG; op deze scholen zitten 4278 leerlingen.

Teneinde te bepalen in hoeverre de steekproef representatief mag worden geacht voor de populatie is de samenstelling van beide op een aantal relevante kenmerken vergeleken. Dit gebeurde zowel op school- als ook op leerlingniveau (vgl. Driessen & Van der Werf, 1992a). De conclusie die op basis van de uitgevoerde analyses kan worden getrokken is dat de steekproef in zijn algemeenheid een redelijke afspiegeling vormt van de populatie. Een uitzondering betreft echter drie aspecten, namelijk: de steden Amsterdam en Den Haag zijn ondervertegenwoordigd, evenals ibo-scholen en -leerlingen; waarschijnlijk daarmee samenhangend is er tevens sprake van een lichte ondervertegenwoordiging van leerlingen in achterstandssituaties.

Voor de hierna te beschrijven analyses is gebruik gemaakt van de volgende variabelen:

Bij de bepaling van de *etnische herkomst* van de leerlingen is er van uitgegaan dat wanneer één van de ouders in het buitenland is geboren, hun kinderen als allochtoon worden beschouwd. De volgende categorieën zijn onderscheiden: (1) Nederlanders, (2) Antillianen, Surinamers en Molukkers, (3) Marokkanen, (4) Turken, (5) overige OVB-groepen (d.w.z. allochtonen die qua etnische herkomst tot de OVB-doelgroepen behoren, bv. Spanjaarden, Grieken), (6) overige allochtonen (bv. Fransen, Engelsen). Deze categorisering heeft plaats gevonden op basis van getalsmatige en inhoudelijke criteria.

Voor de vaststelling van het *beroepsniveau* is eerst voor de vader en de moeder afzonderlijk nagegaan tot welke beroepsgroep zij behoorden. Dat is gebeurd aan de hand van de CBS-beroepenclassificatie. Vervolgens is het hoogste niveau binnen het gezin genomen. De gehanteerde categorieën zijn (grofweg aangeduid): (1) niet-werkzame personen, (2) arbeiders, (3) bedrijfshoofden in de landbouw, (4) lagere employés en bedrijfshoofden buiten de landbouw, (5) middelbare employés, (6) hogere employés en beoefenaren van hogere beroepen.

Het *opleidingsniveau* van de ouders is op

vergelijkbare wijze bepaald als het beroepsniveau. Daarbij is gebruik gemaakt van de Standaard Onderwijsindeling (SOI) van het CBS. Uiteindelijk is het hoogst voltooide niveau binnen het gezin genomen. De indeling is als volgt: (1) geen lager onderwijs voltooid, (2) lager onderwijs, (3) voortgezet onderwijs, lagere trap (bv. lbo, mulo, lagere leerjaren hbs), (4) voortgezet onderwijs, hogere trap (bv. hogere leerjaren hbs, mbo), (5) hoger onderwijs, eerste fase (bv. hbo), (6) hoger onderwijs, tweede fase (bv. universiteit).

In het basisonderwijs wordt de eventuele achterstand van de leerlingen uitgedrukt in het zogenaamde *OVB-gewicht*. Daarbij worden in het algemeen drie categorieën onderscheiden, namelijk leerlingen zonder achterstanden (gewicht 1.00), autochtone achterstandsleerlingen[1] (gewicht 1.25) en allochtone leerlingen in een achterstandssituatie (gewicht 1.90). Hoewel het in het onderhavige onderzoek om het voortgezet onderwijs gaat, zijn - teneinde een compacte indicering van achterstanden te krijgen - ook hier de leerlingen ingedeeld in de drie genoemde categorieën. Daarvoor is gebruik gemaakt van de informatie over de gezinssamenstelling, de etnische herkomst en het beroeps- en opleidingsniveau van de ouders.

Het *onderwijsniveau* betreft het type voortgezet onderwijs waarin de leerlingen zich ten tijde van de start van het onderzoek bevonden.

Voor het vaststellen van de *toetsprestaties* is een speciale CITO-Entreetoets ontwikkeld. Deze kent drie onderdelen, te weten: taal, rekenen en informatieverwerking, met elk 20 items. Naast een score per subtoets is tevens de totaalscore berekend. Deze scores zijn vervolgens omgezet in scores met een gemiddelde van 50 en een standaarddeviatie van 10.

OVG-participatie. Van alle leerlingen is bekend of zij op een school zitten die deelneemt aan een OVG.

2 Sociaal-etnische achtergronden

In Tabel 1 worden de leerlingen buiten en binnen OVG's vergeleken op enkele achtergrondkenmerken, te weten: etnische herkomst, beroeps- en opleidingsniveau ouders[2].

Tabel 1

Etnische herkomst (procenten), respectievelijk beroeps- en opleidingsniveau ouders, naar OVG-participatie (gemiddelden en ( ) standaarddeviaties)

	niet-OVG	OVG	totaal	*N*
Nederland	90.4	77.0	87.6	
Ant/Sur/Mol.	1.7	4.2	2.2	
Marokko	1.3	4.7	2.0	
Turkije	1.1	6.0	2.1	
Overig OVB	.7	1.5	.9	
Overig all.	4.9	6.5	5.3	
N	12985	3371	16356	
beroep	3.7	3.0	3.5	16729
	(1.7)	(1.7)	(1.7)	
opleiding	3.7	3.2	3.6	16710
	(1.1)	(1.2)	(1.1)	

Deze tabel laat goed het grote verschil zien in sociaal-etnische samenstelling van de leerlingenpopulatie binnen en buiten gebieden. Met name de verdeling van de allochtone leerlingen trekt de aandacht. Bijna een kwart van de leerlingen binnen gebieden is van allochtone herkomst; buiten de gebieden gaat het om minder dan 10 procent. Met betrekking tot de verdeling naar beroeps- en opleidingsniveau ouders bedraagt het verschil tussen gebieden en niet-gebieden bijna een halve standaarddeviatie. Overigens behoren lang niet alle allochtone leerlingen tot de doelgroep van het OVB; een aantal van hen voldoet namelijk niet aan alle criteria die voor hen gelden. Dat wil dus zeggen dat het opleidings- en/of beroepsniveau van hun ouders als zijnde te hoog wordt beschouwd. Om een betere indruk te verkrijgen van de verdeling van de OVB-doelgroepen binnen en buiten OVG's, wordt in Tabel 2 de verdeling gegeven van de drie grootste categorieën.

Tabel 2

OVB-gewicht, naar OVG-participatie (procenten)

	niet-OVG	OVG	totaal
1.00	63.8	44.7	59.9
1.25	32.3	39.9	33.8
1.90	4.0	15.4	6.3
N	12911	3342	16253

Op de eerste plaats wordt uit deze tabel duidelijk dat het totaal aantal leerlingen in achterstandssituaties wel erg groot is: meer dan 40 procent van de kinderen uit de steekproef bevindt zich volgens de huidige normen in een

achterstandssituatie. Aangezien in de steekproef enkele grote steden ondervertegenwoordigd zijn houdt dit in dat het genoemde percentage waarschijnlijk nog een onderschatting betreft. Tesser e.a. (1991) komen voor het basisonderwijs op een nog hoger percentage uit, namelijk 48[3]. De vraag kan worden gesteld of deze verhouding met betrekking tot het in achterstandssituatie verkeren wel reëel is.

Opvallend is verder dat de verschillen tussen OVG's en niet-OVG's qua relatieve aantallen 1.25-leerlingen minder groot zijn dan die qua 1.00- en 1.90-leerlingen. Dit heeft waarschijnlijk te maken met het gegeven dat de gebieden vaak samenvallen met (wijken binnen) grote steden (vgl. Tesser e.a., 1991) en dat juist in die steden de concentratie van 1.90-leerlingen het hoogste is. In totaal bedraagt het percentage doelgroepleerlingen buiten de gebieden ruim 36 en binnen de gebieden 55; binnen de OVG's gaat het dus om de helft meer 1.25-plus 1.90-leerlingen dan buiten de OVG's. Dit laatste is een kwantitatieve interpretatie; in kwalitatief opzicht kan men zich bovendien afvragen of er ook niet nog eens sprake is van een diepgaandere en meerzijdige achterstandsproblematiek binnen gebieden.

Tabel 3
Onderwijsniveau, naar OVG-participatie (procenten)

	niet-OVG	OVG	totaal
ibo	2.9	7.5	3.9
lbo	19.8	33.3	22.8
lbo/mavo	6.6	5.4	6.3
mavo	25.3	13.8	22.8
avo/vwo	16.4	26.3	18.6
havo/vwo	23.1	1.0	18.2
vwo	3.5	-	2.7
overig	2.4	12.7	4.7
N	15246	4278	19524

3 Onderwijsniveau

Tabel 3 geeft de verdeling weer van de leerlingen over de onderwijsniveaus binnen en buiten voorrangsgebieden. Ter verduidelijking: 'avo/vwo' omvat de brede brugklassen mavo/havo/vwo; onder 'overig' vallen met name de internationale schakelklassen (isk) en middenschool.

Uit Tabel 3 kan worden afgeleid dat het vooral de lagere onderwijstypen zijn die aan OVG's participeren. Binnen de OVG-categorie bevinden zich vooral veel meer ibo- en lbo-leerlingen en nagenoeg geen havo/vwo- en zelfs helemaal geen vwo-leerlingen. Concreter: van de leerlingen binnen de gebieden volgt ruim 40 procent een ibo- of lbo-opleiding, terwijl het overeenkomstige percentage buiten de gebieden slechts de helft daarvan bedraagt. Uiterst schril steken de percentages leerlingen van de onderscheiden groepen op havo/vwo of vwo tegen elkaar af: 26.6 procent buiten tegen niet meer dan 1 procent binnen de gebieden. Opvallend is dat de OVG-categorie wel een oververtegenwoordiging kent van leerlingen uit brede, heterogene brugklassen van het type mavo/havo/vwo. Misschien heeft dit te maken met het feit dat het hier om een wat vage schoolkeuze betreft, waarbij de leerlingen nog vele kanten op kunnen. Nadere analyse maakt duidelijk dat relatief veel 1.90-leerlingen dit schooltype kiezen; mogelijk hebben zij (of hun docenten) meer moeite met het bepalen van hun definitieve keuze.

Teneinde een duidelijker beeld te krijgen van de verdeling van de achterstandscategorieën over de onderwijsniveaus zijn deze in Tabel 4 nog eens uitgesplitst naar OVG-participatie. Ter wille van de leesbaarheid zijn de percentages afgerond op hele getallen.

Tabel 4
OVB-gewicht naar OVG-participatie, naar onderwijsniveau (procenten)

niveau	ibo		lbo		lbo/mavo		mavo		avo/vwo		havo/vwo		vwo	
OVG	nee	ja	nee	ja	nee	ja	nee	ja	nee	ja	nee	ja	nee	ja
1.00	32	26	41	31	47	35	61	53	75	58	83	79	95	-
1.25	58	51	52	57	49	48	34	34	22	24	15	18	4	-
1.90	10	24	7	12	4	17	5	13	3	18	1	3	1	-
totaal	63	37	70	30	84	16	87	13	70	30	99	1	100	-
N	557		3641		985		3904		3043		2969		433	

Voor zowel 1.00-leerlingen binnen als 1.00-leerlingen buiten OVG's geldt dat hun relatieve aantal sterk toeneemt naarmate het onderwijsniveau stijgt. Voor nagenoeg alle niveaus gaat het omgekeerde op voor 1.25-leerlingen. Het breekpunt lijkt voor beide categorieën te liggen tussen lbo/mavo- en mavo-niveau, of anders gezegd: tussen beroepsonderwijs enerzijds en avo en vwo anderzijds. Het relatieve aantal 1.90-leerlingen vertoont een dalende lijn voor niet-OVG's; voor OVG's is er geen duidelijke trend waarneembaar. Indicatief voor de onderwijspositie van allochtone doelgroepleerlingen is wellicht het vermelde percentage voor vwo: slechts 1 procent van hen bezoekt dit schooltype; het vwo wordt bijna geheel bevolkt door leerlingen uit de 1.00-categorie, dat wil zeggen de middelbare en hogere milieus.

4 Toetsprestaties

Eerder is gebleken dat er grote verschillen bestaan tussen OVG's en niet-OVG's qua sociaal-etnische samenstelling van hun leerlingenpopulaties. Te verwachten valt dat de verschillen in toetsprestaties naar OVG-participatie navenant zullen zijn. In hoeverre dat werkelijk zo is, staat in Tabel 5. Pro memorie: het totaalgemiddelde van de toets en van de subtoetsen bedraagt 50 met een standaarddeviatie van 10.

Tabel 5
Toetsprestaties, naar OVG-participatie (gemiddelden)

	niet-OVG	OVG	*N*	E^2
totaal	51.2	45.5	18159	.05
taal	51.2	45.8	18526	.05
rekenen	51.1	46.2	18456	.05
info	51.1	46.1	18335	.04

Inderdaad laten de gegevens uit Tabel 5 zien dat er redelijke verschillen bestaan tussen leerlingen binnen en buiten onderwijsvoorrangsgebieden; zij bedragen voor de toets totaal en voor elk van de subtoetsen ongeveer een halve standaarddeviatie.

Omdat in de vorige tabellen is gebleken dat niet alle onderwijsniveaus gelijk zijn verdeeld over gebieden en niet-gebieden is de vergelijking die in Tabel 5 wordt gemaakt niet helemaal opportuun. Immers, het ligt voor de hand dat in hogere onderwijsniveaus hogere prestaties worden behaald dan in lagere niveaus. De leerlingen zijn vanuit het basisonderwijs doorgestroomd naar de verschillende typen van voortgezet onderwijs juist vanwege hun verschillen in presteren. Een correctie voor onderwijsniveau is daarom noodzakelijk. In Tabel 6 volgt derhalve per onderwijsniveau een overzicht van de toetsprestaties binnen en buiten gebieden. Onderin de tabel is voor de toets als geheel nog eens apart het gemiddelde per onderwijsniveau vermeld. Opgemerkt dient te worden dat de gemiddelden die zijn berekend voor leerlingen binnen OVG's in onderwijsniveau havo/vwo betrekking hebben op relatief geringe aantallen, namelijk 41. In het vwo bevinden zich helemaal geen leerlingen van scholen die participeren aan een OVG.

Op de eerste plaats wordt uit Tabel 6 duidelijk dat de gemiddelde toetsprestaties inderdaad flink stijgen per onderwijsniveau; het gaat vrijwel steeds om ongeveer een halve standaarddeviatie per niveau. Daarbij is er echter een uitzondering: het verschil tussen ibo- en lbo-leerlingen is aanzienlijk groter dan tussen de andere niveaus. Tussen ibo en lbo bevindt zich een duidelijk breekpunt. Op de tweede plaats blijkt dat de verschillen tussen leerlingen op scholen binnen en buiten gebieden minimaal zijn. Wel is het zo dat binnen OVG's heel

Tabel 6
Toetsprestaties naar OVG-participatie, naar onderwijsniveau (procenten)

niveau	ibo		lbo		lbo/mavo		mavo		avo/vwo		havo/vwo		vwo	
OVG	nee	ja	nee	ja	nee	ja	nee	ja	nee	ja	nee	ja	nee	ja
totaal	35	34	43	42	45	43	49	47	54	53	60	61	64	–
taal	37	35	43	42	47	45	50	48	54	53	59	59	62	–
rekenen	37	36	44	43	45	44	49	47	54	53	60	60	63	–
info	37	36	44	43	46	43	49	48	54	52	59	59	62	–
totaal	35		43		45		49		54		60		64	
N	618		4203		1116		4297		3351		3289		474	

consequent wat lager wordt gepresteerd dan daarbuiten; in het algemeen gaat het om 1 punt verschil. Dit laatste geldt binnen elk van de niveaus en voor elk van de onderscheiden toetsonderdelen.

Aangezien de sociaal-etnische samenstelling van de leerlingenpopulaties binnen en buiten gebieden nogal verschilt, is - behalve voor onderwijsniveau - ook daar een correctie voor op zijn plaats[4]. Concreet is de vraag aan de orde in hoeverre de in Tabel 6 gesignaleerde verschillen in toetsprestaties daadwerkelijk samenhangen met het feit dat sommige leerlingen op scholen zitten die participeren aan een OVG en andere niet, of dat andere factoren - i.c. verschillen in onderwijsniveau en in sociaal-etnische achtergrond - daarbij een zwaarder gewicht in de schaal leggen. Om dit te achterhalen is variantie-analyse uitgevoerd met toetsprestaties als te verklaren variabele. In de vergelijking is als eerste het onderwijsniveau ingevoerd, daarna de individuele achterstandskenmerken op het niveau van het gezin (i.c. het OVB-gewicht) en als laatste de geografisch bepaalde achterstandskenmerken op het niveau van de school (i.c. OVG-participatie). In de eerste drie kolommen van Tabel 7 staan de gestandaardiseerde regressiegewichten (bèta's); in de laatste twee kolommen staan de toetsscores binnen en buiten gebieden nadat er is gecontroleerd voor de invloed van de andere twee onafhankelijke variabelen.

Tabel 7 laat zien dat bij de verklaring van verschillen in toetsscores het onderwijsniveau de belangrijkste factor vormt. Het belang van de sociaal-etnische achtergronden - geïndiceerd via het OVB-gewicht - is relatief beperkt; het vaak hiervoor gehanteerde relevantiecriterium van bèta ≥.15 wordt niet bereikt. Dit geldt in nog sterkere mate voor OVG-participatie; de hoogste bèta daarvan bedraagt niet meer dan .05. De tabel geeft tevens aan dat, indien wordt gecontroleerd voor het onderwijsniveau en voor individueel bepaalde sociaal-etnische achterstandskenmerken, de verschillen in toetsprestaties naar OVG-participatie nog slechts zeer gering zijn. Voor elk van de drie subtoetsen geldt dat leerlingen binnen voorrangsgebieden nog ongeveer 1 punt lager scoren dan leerlingen buiten voorrangsgebieden. Met andere woorden: als de toetsprestaties van leerlingen in dezelfde onderwijsniveaus en met dezelfde - op individueel niveau gemeten - sociaal-etnische achtergronden binnen en buiten OVG's met elkaar worden vergeleken, dan blijken er slechts nog marginale verschillen te bestaan[5]. De 'cumulatie van achterstanden' die te maken hebben met het woonachtig zijn binnen een OVG[6] keert dus niet terug in de toetsprestaties[7]. Of dat een verdienste is van het feit dat de school participeert aan een OVG, daarover valt vooralsnog niets te zeggen.

5 Sociaal-economische of etnische determinanten?

Sedert enige jaren wordt er een discussie gevoerd naar aanleiding van de vraag of de achterstanden van allochtonen in het onderwijs moeten worden toegeschreven aan hun etnische herkomst of aan dezelfde sociaal-economische kenmerken die ook bij autochtone achterstandsleerlingen verantwoordelijk worden geacht voor hun lage onderwijspositie. Concreet is de vraag aan de orde of de onderwijspositie en -prestaties van allochtone kinderen verschillen van die van autochtone kinderen van ouders met eenzelfde opleidings- en beroepsniveau. Om die vraag te beantwoorden zijn verschillende analyse-technieken gehanteerd die overigens nagenoeg alle tot identieke resultaten leiden. Op een enkele uitzondering

Tabel 7

Het effect van onderwijsniveau, OVB-gewicht en OVG-participatie op toetsprestaties (bèta's) en toetsscores binnen en buiten gebieden na controle voor onderwijsniveau en OVB-gewicht (gemiddelden)

	ond.-niveau	OVB-gewicht	OVG-part.	niet-OVG	OVG
totaal	.68	.13	.05	50.7	49.6
taal	.62	.10	.04	50.7	49.6
rekenen	.61	.09	.03	50.7	49.8
info	.57	.14	.04	50.6	49.5

Tabel 8

Resultaten van regressie-analyse met onderwijsniveau en toetsprestaties als afhankelijke en (a) sociale-achtergrondkenmerken en (b) etnische herkomst als onafhankelijke variabelen, naar OVG-participatie. Bèta's en percentage door (a), resp. additioneel door (b) verklaarde variantie

	OVG	ber.	opl.	ASM	Mar.	Trk.	ov. OVB	ov. all.	% R^2	add. % R^2
ond.niveau	nee	.17	.34	.02	.01	.01	.02	.04	20.9	.2
	ja	.14	.29	.14	.08	.12	.07	.08	10.5	3.9
prestaties	nee	.11	.30	–.03	–.03	–.03	–.01	–.01	15.3	.3
	ja	.06	.31	–.03	–.05	–.05	.01	.01	14.6	.5

na wijzen de tot op heden uitgevoerde analyses uit dat wanneer wordt gecontroleerd voor sociaal-economische achtergronden van de leerlingen er geen of nauwelijks verschillen meer zijn naar etnische herkomst. Met andere woorden, allochtone en autochtone leerlingen halen dan nagenoeg dezelfde prestaties. (Voor een overzicht van het betreffende onderzoek, zie Driessen, 1990a, 1990b, 1993.)

Omdat de tot nu toe uitgevoerde analyses bijna steeds beperkt zijn gebleven tot het basisonderwijs is het van belang na te gaan of de hierboven weergegeven conclusie ook opgaat voor het voortgezet onderwijs. Bovendien is het nu mogelijk een verfijning aan te brengen in de analyses die voordien nog niet is toegepast; bedoeld is hier een onderscheid naar OVG's en niet-OVG's. Wat dit laatste betreft is enige uitleg noodzakelijk. Tot nu toe werd in analyses impliciet aangenomen dat de variabele 'etnische herkomst' min of meer normaal verdeeld was over de populatie. Dit is echter in werkelijkheid natuurlijk niet het geval; allochtonen zijn namelijk sterk geconcentreerd in grote steden, terwijl zij daarbuiten slechts in relatief geringe aantallen woonachtig zijn. De op basis van dit type analyse verkregen resultaten zijn wellicht in theoretisch opzicht wel correct maar hebben betrekking op een artificiële, niet bestaande situatie; het zijn eigenlijk louter theoretische 'gemiddelden'. Het benadert meer de werkelijkheid en is daardoor wellicht adequater wanneer de analyses apart worden uitgevoerd voor regio's met veel en regio's met weinig of geen allochtonen. De in het onderhavige onderzoek gehanteerde indeling naar OVG-participatie biedt in deze goede mogelijkheden; de gebieden vallen immers vrijwel steeds samen met (delen van) grote steden.

Rekening houdend met de hierboven gemaakte opmerkingen zijn er twee analyses uitgevoerd; één voor de groep leerlingen op scholen die participeren aan een OVG en één voor de groep leerlingen op scholen die niet samenwerken in het kader van een OVG. Als analyse-techniek is gekozen voor stapsgewijze multiple regressie-analyse. De afhankelijke variabelen zijn eerst het onderwijsniveau en vervolgens de toetsprestaties (totaal). Als onafhankelijke variabelen is als eerste stap in één blok beroeps- en opleidingsniveau ouders ingevoerd – als indicatoren voor de sociaal-economische achtergrondkenmerken – en is daar vervolgens in de tweede stap etnische herkomst aan toegevoegd. Voor dit laatste is de oorspronkelijke variabele uiteengelegd in zes zogenaamde dummies. Vijf daarvan zijn in de tweede stap opgenomen; de zesde – de groep Nederlandse leerlingen – fungeert als referentiecategorie. In Tabel 8 staan de resultaten van de analyses. Van links naar rechts worden de volgende cijfers vermeld: in kolom 1 – 7 de gestandaardiseerde regressiecoëfficiënten (bèta's) voor elk van de onafhankelijke variabelen; in kolom 8 het percentage variantie dat wordt 'verklaard' door beroep en opleiding ouders; in kolom 9 het percentage variantie dat extra wordt 'verklaard' nadat het effect van beroep en opleiding al is verdisconteerd. ('ASM' staat voor de Antilliaanse, Surinaamse en Molukse leerlingen.)

Wat het onderwijsniveau betreft, blijkt uit Tabel 8 dat voor de verklaring van verschillen daarin het opleidingsniveau van de ouders veruit het grootste gewicht heeft. Dit geldt voor zowel leerlingen op scholen binnen als leerlingen op scholen buiten OVG's. Buiten OVG's volgt daarna het beroepsniveau als belangrijkste variabele; de bèta's voor elk van de onderscheiden herkomstland-dummies zijn alle niet relevant. Dit keert ook terug bij de vergelijking van de percentages verklaarde variantie. Bui-

ten gebieden verklaren de sociale-achtergrondkenmerken bijna 21 procent, terwijl etnische herkomst daar minder dan een half procent aan toevoegt. Binnen OVG's zijn de bèta's voor de herkomstlanden hoger dan buiten OVG's en ook het additioneel door etnische herkomst verklaarde percentage variantie is groter dan buiten de gebieden. Toch neemt dit niet weg dat de sociale-achtergrondkenmerken veel belangrijker zijn dan etnische herkomst. Voor de verklaring van verschillen in prestaties zowel binnen als buiten OVG's geldt eigenlijk hetzelfde als wat hierboven is gezegd met betrekking tot onderwijsniveau. Samenvattend lijkt het er dus op dat voor de verklaring van verschillen in onderwijsniveaus en prestaties het opleidingsniveau van de ouders verreweg het grootste gewicht in de schaal legt. Na verdiscontering van de sociale-achtergrondkenmerken voegt de etnische herkomst nagenoeg niets meer toe in deze verklaring[8]. Dit geldt voor de groep leerlingen binnen èn voor de groep leerlingen buiten de gebieden. Eén uitzondering moet worden gemaakt op dit algemene beeld: voor de verklaring van verschillen in onderwijsniveaus binnen de OVG's is er wèl enige - zij het relatief geringe - invloed van etnische herkomst.

6 Conclusies

Er bestaan grote verschillen in sociaal-etnische achtergronden van leerlingen binnen en leerlingen buiten onderwijsvoorrangsgebieden. OVG's kenmerken zich vooral door veel hogere percentages allochtone doelgroepleerlingen en veel lagere percentages leerlingen die niet tot de OVB-doelgroepen behoren. Opvallend is overigens dat volgens de huidige criteria van het OVB ruim 40 procent van de leerlingen als achterstandsleerlingen worden beschouwd. Het is de vraag of dit wel terecht is en of een aanscherping van de regelgeving niet gepast is - dit laatste uiteraard met een herverdeling van de daarmee vrijkomende gelden over de 'echte' achterstandscategorieën.

Een vergelijking van de OVG's naar schooltype laat zien dat het vooral de lagere onderwijsniveaus - en dan met name het ibo en lbo - zijn die aan een OVG participeren; de hogere niveaus (havo/vwo en vwo) komen nagenoeg niet voor binnen een OVG. In het algemeen bestaan er niet-onaanzienlijke verschillen in toetsprestaties tussen leerlingen op scholen binnen en leerlingen buiten gebieden. Nadere analyse maakt echter duidelijk dat deze grotendeels verdwijnen als rekening wordt gehouden met het verschil in onderwijsniveaus tussen OVG's en niet-OVG's. Wanneer vervolgens ook nog het verschil in sociaal-etnische samenstelling van de leerlingenpopulaties wordt verdisconteerd dan blijft er niet meer over dan een verschil van 1 toetspunt bij een gemiddelde van 50. Het voorafgaande samenvattend: verschillen tussen OVG's en niet-OVG's betreffen verschillen in toetsprestaties die nagenoeg geheel kunnen worden herleid tot verschillen in onderwijsniveau (ofwel schooltype) en sociaal-etnische achtergronden op gezinsniveau. Wat het effect van OVG-participatie en de in het kader daarvan de ondernomen activiteiten zal zijn op de prestaties, daarop zal mettertijd pas meer zicht worden gekregen.

Analyses waarmee het relatieve gewicht van sociale achtergronden en etnische herkomst is bepaald bij de verklaring van verschillen in onderwijsniveau en prestaties maken duidelijk dat OVG's daarin nauwelijks verschillen van niet-OVG's. Als rekening wordt gehouden met de sociale achtergronden van de leerlingen (i.c. het opleidings- en beroepsniveau van hun ouders) dan verdwijnen de verschillen tussen allochtone en autochtone leerlingen bijna geheel. Alleen binnen gebieden is er enige extra invloed van etnische herkomst op onderwijsniveau. Dit resultaat wil overigens niet zeggen dat de achterstanden van de 1.90-leerlingen niets van doen hebben met hun etnische herkomst. Immers, door de concentratie van een zeer groot deel van hun ouders in de laagste opleidings- en beroepsniveaus kan hier met evenveel recht worden gesproken van een etnisch als van een sociaal verschijnsel.

Noten

1 Deze categorie wordt ook wel aangeduid als 'Nederlandse arbeiderskinderen'. Helemaal correct zijn beide benamingen niet. Het betreffen namelijk niet allemaal kinderen van arbeiders; bovendien bevat de categorie ook allochtone kinderen die niet tot de OVB-doelgroepen behoren, zoals

Fransen, Duitsers, Engelsen, etc.

2 De verschillen tussen de in de tabellen en in paragraaf 2 genoemde aantallen leerlingen zijn het gevolg van 'missing values'. Analyses met betrekking tot de meeste relevante kenmerken wijzen overigens niet uit dat de non-respons selectief zou zijn (vgl. Driessen & Van der Werf, 1992a).

3 Waarschijnlijk ligt het precieze percentage ergens tussen de 40 en 48 in. VOCL'89 kent als nadeel dat scholen in de Randstad wat ondervertegenwoordigd zijn (m.n. Amsterdam). Een voordeel van VOCL'89 ten opzichte van de LEO-onderzoeken (Landelijke Evaluatie OVB) is echter dat de gegevens over de gezinsachtergronden (o.a. opleiding en etniciteit) van de ouders zelf verkregen zijn en niet via de scholen. In de LEO-onderzoeken in het basisonderwijs zijn de scholen en ouders direct belanghebbenden; de categorisering heeft immers gevolgen voor de omvang van het team en de aandacht voor de kinderen. In het voortgezet onderwijs, waar VOCL'89 betrekking op heeft, is de gewichtenregeling niet van toepassing. Verder zijn de gegevens over opleiding, beroep en etniciteit binnen VOCL'89 door de onderzoekers op uniforme wijze gecategoriseerd, terwijl dat binnen de LEO-onderzoeken aan de afzonderlijke scholen is overgelaten.

4 Overigens wordt in Tabel 8 aandacht besteed aan een analyse waarbij niet is gecontroleerd voor onderwijsniveau.

5 Vanwege de thematiek van deze publikatie is ter indicatie van de individuele achterstandskenmerken het OVB-gewicht gehanteerd. In onderzoek wordt ook wel gebruik gemaakt van een combinatie van het opleidings- en beroepsniveau van beide ouders. Ter vergelijking is nagegaan of dit tot afwijkende resultaten zou leiden. Daartoe is het opleidings- en beroepsniveau van de vader en moeder (voor zover aanwezig) gemiddeld en is vervolgens de nieuwe variabele in vier klassen van elk ca. 25 procent opgedeeld. Deze indicator van het sociaal milieu is daarna ter vervanging van het OVB-gewicht in de analyses opgenomen. Voor de 'toets totaal' zijn de bèta's: .67, .10 en .05, en de gecorrigeerde toetsscores: 50.7 en 49.3. Het beeld voor de drie afzonderlijke toetsonderdelen ligt in dezelfde lijn. De verschillen zijn dus uiterst gering en worden hoogstwaarschijnlijk veroorzaakt doordat bij de milieuvariabele de etnische herkomst sec niet is verdisconteerd. Gezien het bèta-verschil van slechts enkele honderdste punten is dit er tegelijkertijd een indicatie voor dat de etnische herkomst hooguit een zeer minimale rol kan spelen bij de verklaring van verschillen in toetsprestaties.

6 In verband met de situatie rond witte en zwarte scholen zal er sprake zijn van 'grensoverschrijdend verkeer', d.w.z. witte leerlingen zullen wellicht ook naar witte scholen gaan buiten 'hun' gebied.

7 Het is vervolgens de vraag of het hier om een kwalitatief of louter om een kwantitatief probleem gaat.

8 Vergelijk ook het slot van noot 5.

Literatuur

Driessen, G. (1990a). *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. (1990b). Sociaal milieu en etnische herkomst als verklaring voor verschillen in onderwijspositie. *Pedagogisch Tijdschrift, 15,* 349-357.

Driessen, G. (1993). Socio-economic or ethnic determinants of educational opportunities? Results from the educational priority policy programme in the Netherlands. Verschijnt in: *Studies in Educational Evaluation.*

Driessen, G., & Werf, G. van der (1992a). *Het functioneren van het voortgezet onderwijs. Beschrijving steekproef en psychometrische kwaliteit instrumenten.* Groningen/Nijmegen: RION/ITS.

Driessen, G., & Werf, G. van der (1992b). *Het functioneren van het voortgezet onderwijs. De positie van leerlingen in het eerste leerjaar.* Nijmegen/Groningen: ITS/RION.

Kloprogge, J. (red.) (1989). *Voorrang voor de gebieden? Notitie over het onderwijsvoorrangsbeleid in 1988.* 's-Gravenhage: SVO.

Kloprogge, J. (1991). *Het gewicht gewogen. Notitie over het onderwijsvoorrangsbeleid in 1990.* 's-Gravenhage: SVO.

Ministerie van Onderwijs en Wetenschappen (1985a). *Onderwijsvoorrangsplan.* 's-Gravenhage: Staatsuitgeverij.

Ministerie van Onderwijs en Wetenschappen (1985b). *Circulaire BO/OVB-703.218. Aanwijzing onderwijsvoorrangsgebieden per 1 augustus 1986.* Ministerie van Onderwijs en Wetenschappen.

Research voor Beleid (1987a). *Onderwijs en welzijn in onderwijsvoorrangsgebieden.* Leiden: RvB.

Research voor Beleid (1987b). *Onderwijsvoorrangsgebieden, organisatie en management.* Leiden: RvB.

Stimulans (1985). De ontwerp-circulaire. Nieuwsbrief, 6e nummer. *Stimulans.*

Tesser, P., Werf, M. van der, Mulder, L., & Weide, M. (1989). *De beginfase van het voorrangsbeleid in het basisonderwijs. Verslag van het oriënterend onderzoek van de landelijke OVB-evaluatie.* Nijmegen/Groningen, ITS/RION.

Tesser, P., Mulder, L., & Werf, M. van der (1991). *De eerste fase van de longitudinale OVB-onderzoeken. Het leerlingenonderzoek.* Nijmegen/Groningen: ITS/RION.

Verhaak, C., & Brandsma, H. (1991). *Onderwijsvoorrangsgebieden op weg. Een beschrijving van de gebieden als organisatienetwerken en projectorganisaties.* Nijmegen/Groningen: ITS/RION.

Vermeulen, H., & Werf, M. van der (1989). *De plannen van de onderwijsvoorrangsgebieden. Verslag van een analyse van de OVB-gebiedsplannen.* Groningen: RION.

Manuscript aanvaard 11-1-1993

Auteur

G. Driessen is werkzaam op het Instituut voor Toegepaste Sociale Wetenschappen te Nijmegen.

Adres: ITS, Postbus 9048, 6500 KJ Nijmegen

Abstract

Differences between pupils in and out of Educational Priority Areas

G. Driessen, Pedagogische Studiën, 1993, 70, 252-262.

One of the components of the Educational Priority Policy concerns the policy regarding the Educational Priority Areas (EPAs). Within EPAs there is an accumulation of educational disadvantages. Primary schools, secondary schools and welfare agencies work together to combat these disadvantages. This article focuses on differences between secondary education pupils in and out of EPAs. In the analyses a sample of nearly 20,000 pupils is used. The results show that there are differences between EPAs and non-EPAs concerning social and ethnic backgrounds of the pupils, level of secondary education and test scores. However, the differences in test scores can be attributed to differences in level of education which in turn can be ascribed to differences in social-ethnic backgrounds.

Samenwerking in Onderwijsvoorrangsgebieden*

C. Verhaak

Samenvatting

Het doel van onderwijsvoorrangsgebieden is om in regio's waar sprake is van een cumulatie van achterstandsfactoren door middel van samenwerking tussen scholen en welzijnsinstellingen de onderwijskansen van kinderen te verbeteren. Onderwijsvoorrangsgebieden moeten vormgeven aan schooloverstijgende bestrijding van onderwijsachterstanden, afgestemd op het lokaal beleid. In dit artikel wordt verslag gedaan van een onderzoek naar de stand van zaken van het inhoudelijk beleid van en de samenwerkingsrelaties binnen de onderwijsvoorrangsgebieden. Uit het onderzoek blijkt dat samenwerking tussen scholen en welzijnsinstellingen slechts moeizaam op gang komt. Dat neemt niet weg dat er binnen onderwijsvoorrangsgebieden een breed palet van activiteiten wordt uitgevoerd. De doelstellingen van de activiteiten zijn echter veelal vaag omschreven. Het ontbreekt aan een uitwerking in operationele termen. Dit heeft met name te maken met de onduidelijkheid die er heerst over de wijze waarop onderwijsachterstanden op lokaal niveau moeten worden bestreden.

Inleiding

Sociale ongelijkheid is hardnekkig, niet alleen in het onderwijs. Jarenlange bestrijding ervan heeft nog steeds niet het beoogde resultaat opgeleverd. Er is onvoldoende duidelijk wat de juiste strategieën zijn om de ongelijkheid te verminderen. Ook is er onduidelijkheid over de instrumenten die hiertoe het best kunnen worden ingezet. De motieven om achterstanden te bestrijden wisselen met de geest van de tijd. Was het in de jaren zeventig de roep om rechtvaardigheid en gelijke kansen die, gesteund door een overheid met geloof in de verzorgingsstaat leidde tot grootscheepse stimuleringsprogramma's, de laatste jaren staan efficiëntie en resultaten centraal. De verzorgingsstaat raakt uit de gratie. De moeilijke taak van de achterstandsbestrijding wordt doorgeschoven naar het lokaal beleid. Samenwerking tussen instellingen op lokaal niveau moet er toe leiden dat beleid doelgericht, dicht bij huis en afgestemd op de plaatselijke situatie kan worden vormgegeven en uitgevoerd. Het hoogtepunt van deze samenwerkingsgedachte zien we in de sociale vernieuwing. Ook het beleid rond onderwijsvoorrangsgebieden past in dit kader.

Samenwerking in onderwijsvoorrangsverband moet leiden tot een vergroting van de afstemming van en de continuïteit in activiteiten die tot doel hebben de onderwijskansen van kinderen in sociaal-economisch zwakke regio's te verbeteren. Men doelt hierbij vooral op afstemming van voorrangsactiviteiten van verschillende typen scholen: basisscholen en scholen voor speciaal onderwijs en basisscholen met scholen voor voortgezet onderwijs. Dit moet de continuïteit van het onderwijs en van de achterstandsbestrijding op leerlingniveau tijdens de hele schoolloopbaan waarborgen.

In dit artikel beantwoorden we de vraag wat er van die samenwerking in onderwijsvoorrangsgebieden terecht gekomen is. Hoe intensief is de samenwerking? Wat is de kwaliteit van de samenwerking en welke activiteiten worden er uitgevoerd? Onder welke voorwaarden levert samenwerking de beste resultaten op? Dit is niet alleen van belang om inzicht te krijgen in de wijze waarop onderwijsvoorrangsbeleid vorm krijgt. Het gaat er ook om lering te trekken uit ervaringen die in onderwijsvoorrangsgebieden zijn opgedaan, juist nu in het kader

* Dit artikel is onder andere geschreven op basis van een onderzoek naar de stand van zaken in onderwijsvoorrangsgebieden dat in het kader van de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid door het ITS en het RION is uitgevoerd.

van de sociale vernieuwing veel nieuwe initiatieven tot samenwerking zijn ontstaan en nog zullen ontstaan. Wanneer ook de voorstellen van de commissie Van Kemenade doorgang zullen vinden zal samenwerking tussen scholen en welzijnsinstellingen op lokaal niveau een steeds grotere rol gaan spelen in beleid om onderwijskansen te verbeteren (Ministerie van Onderwijs en Wetenschappen, 1992). Voor we ingaan op de gang van zaken in onderwijsvoorrangsgebieden geven we eerst nog een kort overzicht van de wettelijke regelingen rond onderwijsvoorrang. Ook beschrijven we een inhoudelijk kader van waaruit processen die zich in onderwijsvoorrangsgebieden afspelen begrepen kunnen worden.

1 Achtergrond wetsvoorstel

In de regio's die als onderwijsvoorrangsgebied zijn aangewezen is sprake van een cumulatie van achterstanden. Dit vergroot de onderwijsachterstand van de leerlingen en maakt de bestrijding ervan moeilijker. Verschillende instellingen werken vaak langs elkaar heen met dezelfde doelgroep en bestrijden deels samenhangende problemen. Het belangrijkste doel van het wetsvoorstel is te komen tot een samenhangend onderwijsvoorrangsbeleid. Die samenhang zou bevorderd worden door scholen en welzijnsinstellingen samen te laten werken bij het verbeteren van onderwijskansen in hun regio. Samenhang houdt in dat de werkwijze van instellingen die met eenzelfde doelgroep te maken hebben op elkaar aansluit. Zo is het bijvoorbeeld aan te bevelen dat het voortgezet onderwijs bij de keuze van de lesmethoden aansluit bij de wijze waarop in het basisonderwijs is gewerkt. Zo moeten aansluitingsproblemen tussen verschillende schooltypen (basisschool-speciaal onderwijs; basisschool-voortgezet onderwijs) worden voorkomen. De inbreng van het welzijnswerk wordt van belang geacht om de discrepantie tussen school- en thuismilieu te overbruggen. De gemeente moet er zorg voor dragen dat het beleid van het gebied afgestemd wordt op de lokale situatie en aansluit bij samenhangend beleid. Dit alles moet er toe leiden dat er ten aanzien van het onderwijs een continuïteit ontstaat in de dienstverlening voor kinderen en ouders in een bepaalde regio.

Ook in het wetswijzigingsvoorstel uit 1990 wordt opnieuw gewezen op het belang van samenwerking. De formulering is echter anders: samenwerking is noodzakelijk vanuit het oogpunt van efficiëntie. Om geïntegreerd beleid, afgestemd op plaatselijke situatie mogelijk te maken worden er weinig regels van bovenaf gesteld. Men probeert zoveel mogelijk de vrijheid van invulling op lokaal niveau te versterken. Hetzelfde zien we binnen de sociale vernieuwing. Ook dit beleid staat in het teken van samenwerking op lokaal niveau. Ook nu moet dat leiden tot doelgerichter en efficiënter beleid dat dicht bij de mensen zelf vorm krijgt. Alleen is het beleidsterrein waar het bij sociale vernieuwing om gaat veel groter. Niet alleen lokaal onderwijs- en welzijnsbeleid zijn er bij betrokken, het gaat ook om volkshuisvesting en arbeidsvoorziening.

Samenwerken, afstemmen, continuïteit aanbrengen, deze termen komen terug in verschillende beleidsnotities. Concrete invullingen worden echter niet gegeven. Dat wordt overgelaten aan gemeenten en andere instellingen op lokaal niveau om 'de aansluiting bij de plaatselijke situatie te garanderen'.

2 Inhoudelijke achtergrond

Wanneer we iets willen zeggen over de effectiviteit en de werking van onderwijsvoorrangsgebieden moet duidelijk zijn vanuit welke inhoudelijke invalshoek deze samenwerkingsverbanden bekeken moeten worden; welk inhoudelijk kader biedt aanknopingspunten om processen die zich in voorrangsgebieden afspelen te verklaren? In het verleden zijn onderwijsstimuleringsprojecten beschouwd als projectorganisaties (Smets & Van Kessel, 1980; Research voor Beleid, 1987a, 1987b). Onderwijsvoorrangsgebieden zijn echter maar ten dele als projectorganisaties te begrijpen.

Een belangrijk kenmerk van een projectorganisatie is dat de projectleiding casu quo het management in handen is van een of meerdere personen met bepaalde (gedelegeerde) bevoegdheden en duidelijk gedefinieerde taken. Binnen de organisatie van de onderwijsvoorrangsgebieden is weliswaar een belangrijke taak weggelegd voor het projectmanagement casu quo de coördinator, de formele bevoegd-

heid van dit management is nihil. Op bestuursniveau hebben de coördinatoren slechts een adviserende stem, naar het uitvoeringsniveau ontbreekt het hen aan sanctionerende bevoegdheden. Deelnemende organisaties zijn primair verantwoording schuldig aan hun eigen bevoegd gezag. Het samenwerkingsverband zelf is geen rechtspersoon en kan ook geen mensen in dienst nemen.

Onderwijsvoorrangsgebieden zijn daarom beter te begrijpen als organisatienetwerken. De samenwerkende instellingen blijven zelf verantwoordelijk voor de inzet van de middelen, hun samenwerking heeft echter wel een min of meer gestructureerde vorm. In het onderwijs is deze netwerkontwikkeling nog tamelijk jong. De autonomie van instellingen, met name van scholen, heeft van oudsher hoog in het vaandel gestaan. Pas de laatste jaren ontstaat er meer samenwerking tussen scholen onderling, tussen scholen en overheidsinstellingen en tussen scholen en andere organisaties.

Het meest opvallende dat in de wetenschappelijke literatuur over organisatienetwerken wordt benadrukt, is hun ambigue karakter. Organisaties zijn bovenal gesteld op hun eigen autonomie. Ze gaan pas relaties met elkaar aan als ze hier voordeel uit kunnen halen. Netwerken bewegen zich daarom steeds in het spanningsveld van autonomie versus afhankelijkheid: "Als we er vanuit gaan dat de meeste organisaties gesteld zijn op hun autonomie en identiteit, dan is aan te nemen dat de meeste netwerken vaker uit nood of expansiedrang dan uit zoiets als liefde-op-het-eerste-gezicht worden geboren" (Wassenberg, 1980, p. 17).

Ook in netwerken van onderwijsinstellingen is deze ambiguïteit te herkennen: "Samenwerkingsverbanden van scholen zijn instabiel omdat iedere school in vrijheid steeds kan afwegen of de baten van het samenwerkingsverband nog opwegen tegen de lasten" (Van der Geest, 1986, p. 3).

Er zal dus altijd worden gezocht naar een manier om de voordelen die het netwerk biedt te vergroten zonder dat dit ten koste gaat van de autonomie van de eigen instelling, van de eigen school. Daarom is het voor een helder inzicht in het functioneren van netwerken niet voldoende om de doelstellingen van het netwerk als geheel te kennen. Iedere organisatie maakt namelijk constant afwegingen tussen het eigen belang en het belang van het netwerk. Belangen van de afzonderlijke instellingen bepalen dan ook minstens zo sterk het functioneren van het netwerk als de doelstellingen van het netwerk in het geheel. Samenwerking in netwerken is vaak ver te zoeken. Deelnemers zijn er juist op uit om de eigen autonomie te allen tijde te bewaken. De eigen belangen van de instellingen komen dan gemakkelijk in conflict met het gemeenschappelijk doel van het samenwerkingsverband. Bij de analyse van netwerken moeten daarom de afzonderlijke organisaties centraal staan: "Zij onderhandelen, concurreren en werken eventueel samen" (Godfroij, 1981). Naast de behoefte om samen te werken moeten ook de tegengestelde belangen duidelijk zijn.

Dit ambigue karakter van organisatienetwerken brengt met zich mee dat de structuur van het netwerk niet zozeer bepaald wordt door de reden van samenwerking maar eerder door de eigen belangen van de organisaties. Er is een constante spanning tussen het belang van het netwerk en het belang van de deelnemers. Dit brengt sommige auteurs (Wassenberg, 1980 en ook Godfroij, 1981) er toe netwerken te beschrijven als arena's. In die arena's spelen participanten constant een spel waarin ieder een evenwicht zoekt tussen autonomie en afhankelijkheid. In de volgende paragraaf wordt duidelijk hoe het spel in de arena van het onderwijsvoorrangsgebied van 1986 tot 1990 is gespeeld.

3 Onderzoeksopzet

Ten behoeve van het onderzoek naar de stand van zaken in onderwijsvoorrangsgebieden zijn in 1988 en 1989 36 gebiedscoördinatoren benaderd. Dat is vijftig procent van het totale aantal coördinatoren dat ten tijde van het onderzoek actief was. Deze coördinatoren waren verbonden aan de OVG's van de steekproef van de cohortstudies van de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid. Hierdoor is het in de toekomst mogelijk geworden gegevens van deze gebieden te koppelen aan gegevens op schoolniveau.

De coördinatoren hebben in het najaar van 1988 een schriftelijke vragenlijst ontvangen waarin vragen werden gesteld over bestuurlijke samenstelling, organisatorische vormge-

ving, interne en externe samenwerkingsrelaties en activiteiten van het gebied. Dertig coördinatoren waren bereid met het onderzoek mee te doen. Bij hen is ook een interview afgenomen. Tijdens dat interview zijn achtergronden van de gegevens uit de schriftelijke vragenlijst besproken.

4 Resultaten: de praktijk in onderwijsvoorrangsgebieden

4.1 Intensiteit van de samenwerking

Participanten in onderwijsvoorrangsgebieden ontmoeten elkaar in de meeste gevallen niet frequent. Op beleidsbepalend niveau (directies) treft men elkaar eigenlijk met name op jaarvergaderingen van het samenwerkingsverband. Alleen directies van basisscholen onderling zitten gemiddeld vaker rond de tafel: men treft elkaar zelfs gemiddeld gesproken maandelijks. Op uitvoerend niveau zijn de contacten meestal wat hechter. Basisscholen en welzijnsinstellingen werken het intensiefst samen (gemiddeld maandelijks). Basisscholen onderling, scholen voor voortgezet onderwijs onderling en scholen voor voortgezet onderwijs en welzijnsinstellingen treffen elkaar op uitvoerend niveau enkele malen per jaar.

4.2 Uitgevoerde activiteiten

In het totaal zijn in zesendertig onderwijsvoorrangsgebieden voor de periode 1986-1990 1443 activiteiten gepland. Dat zijn er gemiddeld ongeveer veertig per gebied. Deze activiteiten zijn onder te verdelen in tien verschillende thema's:

- basisvaardigheden
  activiteiten gericht op het verbeteren van taal- en rekenvaardigheid;
- zorgverbreding
  stimuleren van randvoorwaarden om cognitieve ontwikkeling te verbeteren zoals leerlingvolgsysteem, leerlingbegeleiding en schoolmaatschappelijk werk;
- niet-cognitieve ontwikkeling
  activiteiten ter stimulering van creatieve, sportieve en sociaal-emotionele ontwikkeling;
- school-ouders-buurt
  activiteiten voor ouders;
- gebiedsontwikkeling
  activiteiten die te maken hebben met de instandhouding van het samenwerkingsverband zelf;
- gedragsproblematiek
  activiteiten rond opvang van spijbelaars en preventie van schooluitval;
- roldoorbreking
  activiteiten die tot doel hebben stereotiepe rolpatronen van mannen en vrouwen/jongens en meisjes ter discussie te stellen;
- voorschoolse periode
  activiteiten voor kinderen tot vier jaar en hun ouders;
- intercultureel onderwijs
  activiteiten die het multi-culturele karakter van de samenleving expliciet uitdragen;
- ononderbroken ontwikkelingslijn
  activiteiten die er toe moeten leiden dat de aansluiting tussen verschillende typen onderwijs verbetert.

In Tabel 1 wordt duidelijk hoe de activiteiten over deze tien thema's zijn verdeeld.

Tabel 1
*Activiteiten naar thema (N=1443)**

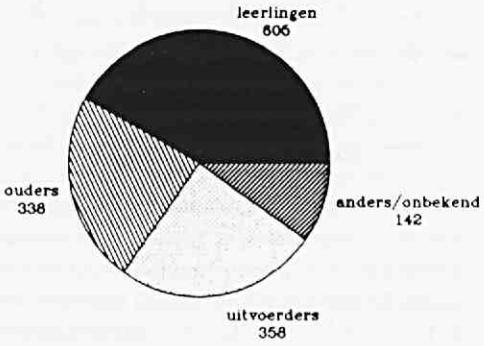
thema	*N*	%
basisvaardigheden	331	23
zorgverbreding	152	11
niet cognitieve ontw.	129	9
school-ouders-buurt	291	20
gebiedsontwikkeling	53	4
gedragsproblematiek	60	4
roldoorbreking	37	2
voorschoolse periode	50	4
intercultureel onderwijs	109	7
ononderbroken ontw.lijn	120	8
anders	111	8
totaal	1443	100

* Tabel uit Verhaak en Brandsma, 1991

Ruim de helft van de voorrangsactiviteiten heeft betrekking op basisvaardigheden, zorgverbreding en het verbeteren van de contacten tussen school-ouders en buurt. Hierbij moet in ogenschouw worden genomen dat het hierbij gaat om informatie uit de gebiedsplannen die in 1986 zijn opgesteld. De laatste jaren is met name het thema 'voorschoolse periode' sterk in de aandacht gekomen. Veel onderwijsvoorrangsgebieden hebben zich op dit terrein ge-

profileerd (Verhaak & Kristensen, 1992). De nadruk op basisvaardigheden zal verder eerder zijn toe- dan zijn afgenomen.

Ruim veertig procent van de activiteiten is direct gericht op de leerlingen zelf. Een kwart van de activiteiten is gericht op ouders, bij nog een kwart gaat het om activiteiten voor leerkrachten. Vaak gaat het hier om bijscholing. In Figuur 1 wordt de verdeling van de activiteiten naar doelgroep schematisch weergegeven.



Figuur 1. Activiteiten naar doelgroep (absolute aantallen, *N*=1443)

In onderwijsvoorrangsgebieden werken basisscholen, scholen voor voortgezet onderwijs, scholen voor speciaal onderwijs en welzijnsinstellingen samen. Per activiteit gaan verschillende typen participanten met elkaar samenwerken. Het kan gaan om een samenwerkingsproject tussen basisscholen onderling, het is ook mogelijk dat basisscholen samen met bijvoorbeeld scholen voor speciaal onderwijs en welzijnsinstellingen bij een activiteit betrokken zijn. In Tabel 2 wordt duidelijk hoe vaak in welke constellatie wordt samengewerkt.

Basisscholen en welzijnsinstellingen zijn bij de meeste activiteiten betrokken. Basisscholen leveren een bijdrage aan 59% van de activiteiten, welzijnsinstellingen aan ruim 55%. Het voortgezet onderwijs is betrokken bij bijna 40% van de activiteiten, het speciaal onderwijs bij nog geen kwart. In dit artikel blijft de rol van het speciaal onderwijs verder buiten beschouwing. Petersen en De Keijzer (1992) hebben over de positie van het speciaal onderwijs binnen onderwijsvoorrang uitgebreid gerapporteerd.

In Tabel 3 wordt aangegeven welke soorten voorrangsactiviteiten door de verschillende samenwerkende instellingen worden uitgevoerd.

Algemeen beschouwd (zie eerste kolom van Tabel 2) ligt het zwaartepunt van activiteiten van onderwijsvoorrangsgebieden rond de thema's basisvaardigheden en school, ouders en buurt. Bijna de helft van de activiteiten heeft een van deze twee thema's. In de rechter zes kolommen van Tabel 3 worden deze gegevens verder uitgesplitst naar type samenwerking.

Tabel 2
*Activiteiten naar samenwerkingsrelatie**

samenwerkingsrelatie	aantal	%
basisscholen onderling (BOBO)	131	9.1
basisscholen met vo-scholen (BOVO)	58	4.0
basisscholen met so-scholen (BOSO)	48	3.3
basisscholen met welz.instellingen (BOWW)	256	17.2
vo-scholen onderling (VOVO)	99	6.9
vo- en so-scholen (VOSO)	13	.9
vo-scholen met welz.instellingen (VOWW)	109	7.6
so-scholen onderling (SOSO)	7	.5
so-scholen met welz.instellingen (SOWW)	8	.6
welzijnsinstellingen onderling (WWWW)	131	9.1
basischolen met so- en vo-scholen (BOVOSO)	71	4.9
basischolen met vo-scholen en welz. (BOVOWW)	94	6.5
basischolen met so-scholen en welz. (BOSOWW)	74	5.1
vo- en so-scholen met welz. (VOSOWW)	11	.8
alle typen participanten samen	115	8.0
onduidelijk	216	15.0
totaal	1443	100.0

* Tabel uit Verhaak en Brandsma, 1991

Tabel 3
*Thema van de activiteit naar samenwerkingsrelatie (in procenten; N=1443)**

thema	tot.	BOBO	WWWW	VOVO	BOVO	BOWW	VOWW
basisvaardigheden	23	**49**	**21**	**25**	**21**	**22**	**10**
zorgverbreding	11	**16**	4	**20**	**21**	3	**26**
niet cognitieve ontw.	9	6	5	7	7	**15**	**15**
school-ouders-buurt	20	**15**	**37**	3	5	**38**	8
gebiedsontwikkeling	4	1	3	—	—	5	2
gedragsproblematiek	4	1	2	**15**	2	—	**19**
roldoorbreking	3	5	5	5	—	1	6
voorschoolse periode	4	1	**18**	—	2	6	—
intercultureel onderwijs	8	7	3	7	5	7	6
ononderbroken ontw.lijn	8	1	2	**17**	**38**	4	8
anders	8	—	1	—	—	—	1
totaal							

* Tabel uit Verhaak en Brandsma, 1991

Dan blijkt dat verschillende samenwerkende instellingen niet dezelfde typen activiteiten uitvoeren. Basisscholen leggen in hun voorrangsactiviteiten een sterke nadruk op de basisvaardigheden. Relatief minder aandacht wordt besteed aan zorgverbreding en oudercontacten. Het welzijnswerk daarentegen concentreert zich op de ouderactiviteiten en de andere activiteiten die de verschillen tussen leefwerelden thuis en op school proberen te overbruggen. Daarnaast besteedt het welzijnswerk zelfstandig nog zo'n twintig procent van de voorrangsactiviteiten aan basisvaardigheden. Vaak gaat het hierbij om leespromotieprojecten van de bibliotheek. Ook activiteiten in de voorschoolse periode worden veel door het welzijnswerk uitgevoerd. Dit percentage zal inmiddels gestegen zijn. In het voortgezet onderwijs concentreren voorrangsactiviteiten zich rond basisvorming, zorgverbreding en verbetering van de aansluiting tussen verschillende schooltypen. Bij zorgverbreding gaat het hierbij met name om leerlingbegeleiding. Ook huiswerkbegeleiding valt onder dit thema. Daarnaast wordt in het voortgezet onderwijs in het kader van OVB aandacht besteed aan preventie van drop-outs en opvang van spijbelaars.

Basischolen en scholen voor voortgezet onderwijs samen richten zich op het thema ononderbroken ontwikkelingslijn. Het gaat hierbij om het verkleinen van aansluitingsproblemen tussen basis- en voortgezet onderwijs.

4.3 Kwaliteit van de samenwerking in cijfers

Aan de coördinatoren is gevraagd per samenwerkingsrelatie op een vijfpuntsschaal (1= zeer slecht, 5= zeer goed) hun oordeel te geven over de kwaliteit. Het resultaat hiervan is te zien in Tabel 4.

Tabel 4
*Kwaliteit van de samenwerkingsrelatie (N=1443)**

type samenw.	*gem.*	*sd*
BOBO	3.8	1.22
BOVO	2.9	0.75
BOSO	3.0	0.96
BOWW	3.7	0.67
VOVO	3.6	0.79
VOSO	3.0	0.83
VOWW	3.2	0.86
SOSO	3.2	0.89
SOWW	3.0	1.09
totaal	3.2	0.57

* Tabel uit Verhaak en Brandsma, 1991

De kwaliteit van de samenwerking tussen basisscholen onderling wordt het meest positief beoordeeld. Ook de samenwerking tussen basisscholen en welzijnsinstellingen en tussen scholen voor voortgezet onderwijs onderling verloopt redelijk goed. Het slechtste gaat de samenwerking tussen basisscholen en scholen voor voortgezet onderwijs.

In de nu volgende subparagraaf wordt een overzicht gegeven van factoren die samenwerkingsrelaties positief dan wel negatief beïnvloeden.

De interviews hebben informatie opgeleverd over de achtergrond van deze kwaliteitsverschillen. Aan de coördinatoren is gevraagd naar factoren die van invloed zijn geweest op de samenwerkingsrelaties. Ook is ingegaan op de rol die de coördinatoren zelf hebben kunnen spelen om die samenwerking te stimuleren. Het is opmerkelijk dat de coördinatoren factoren die het functioneren van het samenwerkingsverband beïnvloeden vaag omschrijven. Vaak weet men nauwelijks aanwijzingen voor beter functioneren te formuleren. Om die reden geeft het volgende overzicht van de gegevens die een kwalitatieve analyse van de interviewgegevens heeft opgeleverd hier en daar een diffuus beeld. Hierdoor wordt echter duidelijk met welk een moeizaam proces het bestrijden van onderwijsachterstanden op schooloverstijgend niveau vaak gepaard gaat.

4.4 Kwaliteit van de samenwerkingsrelaties, achtergronden

Basisscholen onderling
De samenwerking tussen basisscholen onderling verloopt in het algemeen het meest vanzelfsprekend. Veel scholen hebben al enige traditie in samenwerking en kennen elkaar van diverse overlegsituaties. Zo hebben scholen die *dicht bij elkaar liggen* in het verleden vaak al met elkaar te maken gehad. Er zijn bijvoorbeeld afspraken geweest voor gezamenlijke wervingsactiviteiten, planning van vakanties of afspraken over schooltijden. Ook scholen van dezelfde zuil hebben regelmatig contact met elkaar. Samenwerking tussen scholen van verschillende denominatie kan erg moeizaam zijn. Men heeft een (deels) overlappend voedingsgebied en is dus elkaars concurrent.

Uit Tabel 3 blijkt dat basisscholen voornamelijk samenwerken op het terrein van de basisvaardigheden en zorgverbreding. Men realiseert zich welke meerwaarde contacten met andere scholen kan opleveren. Men wisselt ideeën uit over onderwerpen als ETC-integratie, oudercontacten en Nederlands als tweede taal. Ook neemt men materialen van elkaar over.

Een andere belemmering voor samenwerking is de beperkte motivatie die sommige scholen hebben om deel te nemen aan een onderwijsvoorrangsgebied.

Het belangrijkste doel blijkt dan de extra faciliteiten te zijn. De basis voor samenwerking is in dit geval gering.

De motivatie voor OVB kan nog zo groot zijn, als de organisatorische voorwaarden ontbreken komt er weinig van de grond. Veel is bijvoorbeeld afhankelijk van de positie die schoolcontactpersonen van het onderwijsvoorrangsgebied op de scholen innemen. Een sterk bestuur kan ook een belangrijke stimulerende rol spelen. Wanneer het bestuur het belang van het onderwijsvoorrangsbeleid onderstreept kan dat een belangrijke steun in de rug betekenen voor de OVB-leerkrachten op school.

Scholen voor voortgezet onderwijs onderling
De samenwerking tussen scholen voor voortgezet onderwijs komt meestal moeizamer van de grond dan bij basisscholen onderling. Contacten tussen de scholen voor voortgezet onderwijs onderling zijn minder vanzelfsprekend dan contacten tussen basisscholen. De VO-scholen werken autonomer en hebben minder makkelijk directe contacten met buurt- en welzijnsorganisaties. Dat heeft onder meer te maken met het grotere voedingsgebied van scholen voor voortgezet onderwijs. Basisscholen hebben vaak een sterke betrokkenheid bij de buurt waarin ze liggen. In veel gebieden wordt deze buurtgerichte benadering nog benadrukt door een wijkgerichte gebiedsstructuur. Scholen voor voortgezet onderwijs hebben minder affiniteit met 'de buurt' en werken meer op stedelijk of zelfs gemeentegrens overschrijdend niveau.

De *onderlinge concurrentie* is vooral in het voortgezet onderwijs groot. Het recruteren van leerlingen staat bovenaan. De gebiedsfaciliteiten kunnen daardoor makkelijk voor andere doelen worden ingezet. Dit veroorzaakt het belangrijkste probleem dat veel coördinatoren ten aanzien van de samenwerking tussen scholen voor voortgezet onderwijs naar voren brengen. Er worden *boventalligen* ingezet op de uren die door het gebied voor OVB beschikbaar zijn gesteld. Deze boventalligen zijn dus niet geselecteerd op basis van hun eventuele geschiktheid voor het uitvoeren of coördineren van OVB-activiteiten op de school maar op basis van hun rechtspositionele status.

Op scholen voor voortgezet onderwijs lijkt samenwerking met andere scholen op meer or-

ganisatorische problemen te stuiten dan op basisscholen. De scholen hebben intern een minder flexibele organisatiestructuur.

Ook *fusies* vormen een belemmering voor het OVB. Wanneer er in een gebied tegelijkertijd enkele fusies plaatsvinden kan dat de betrokkenheid van het voortgezet onderwijs tot een minimum beperken. In oktober 1992 maakte staatssecretaris Wallage bekend dat een kwart van de scholen voor voortgezet onderwijs in 1993 betrokken zal zijn bij een fusie. In sommige streken van het land gaat het zelfs om vijftig procent van de scholen. Veel scholen die in 1993 niet bij een fusie betrokken zijn hebben deze fusie net achter de rug. Met dit gegeven in het achterhoofd kunnen we vraagtekens zetten bij de mogelijkheden die er zijn om onderwijsvoorrang op een degelijke manier binnen scholen voor voortgezet onderwijs van de grond te krijgen.

Het van de grond krijgen van samenwerking tussen scholen voor voortgezet onderwijs doet een extra beroep op de inzet van de coördinatie van het gebied. De coördinatoren wijzen op het belang van het werken met operationele doelen en een planmatige aanpak. Het gezamenlijk opzetten van een spijbelopvangproject of een werkgroep uitvoering leerplicht zijn voorbeelden van zulke activiteiten die met goede geoperationaliseerde doelen en een systematische planning tot goede samenwerkingsresultaten kunnen leiden. Met name in de 'werkgroepen uitvoering leerplicht' worden positieve ervaringen opgedaan met de samenwerking tussen scholen voor voortgezet onderwijs onderling.

Samenwerking tussen basisscholen en scholen voor voortgezet onderwijs.

In lang niet alle gebieden is een samenwerking tussen VO- en basisscholen van de grond gekomen. In veel gebieden is er weliswaar een poging tot samenwerking gedaan, maar tot grote resultaten heeft dat nog niet geleid.

Een belangrijk thema waar basisscholen en scholen voor voortgezet onderwijs zich gezamenlijk op richten is het verbeteren van de aansluiting tussen de beide schooltypen. Het onderlinge vertrouwen blijkt wel eens beperkt te zijn. Dat kan komen door het verschil in voedingsgebied. Ook hier speelt de verzuiling nogal eens een negatieve rol. Soms zijn de visieverschillen tussen scholen zo groot dat er geen gezamenlijke activiteiten kunnen worden uitgevoerd.

Samenwerking tussen basisscholen en welzijnsinstellingen

Het welzijnswerk is moeilijk onder één noemer te vangen. Er zijn diverse werksoorten betrokken bij het voorrangsbeleid. Het gaat met name om de bibliotheken, het jeugd- en jongerenwerk, het maatschappelijk werk en de peuterspeelzalen.

Wanneer er gevraagd wordt naar positieve ervaringen met de samenwerking tussen onderwijs en welzijn wordt de *bibliotheek* vaak genoemd. Oorzaken die hiervoor worden aangehaald hebben te maken met de schijnbaar beperkte cultuurverschillen die er tussen scholen en bibliotheken overbrugd moeten worden. Doelstellingen van bibliotheekactiviteiten kunnen eenvoudig op uitvoerend niveau worden geformuleerd. Wanneer scholen en welzijnsinstellingen meer op beleidsbepalend niveau proberen samen te werken ontstaan er vaak problemen. Verschillen in prioriteitsstelling en methodiek lijken dan moeilijk op te lossen.

Een algemene oorzaak die genoemd wordt voor problematische samenwerking met het welzijnswerk hangt voor een deel samen met de welzijnspolitiek zoals die in de jaren tachtig door de overheid is gevoerd. Onder andere als gevolg van de bezuinigingen wordt het welzijnswerk geplaagd door veel reorganisaties en veranderingen in de organisatiestructuur. Dat brengt wisselingen in personeel met zich mee wat een zorgvuldig opgebouwd contact met een school gemakkelijk kan verstoren. Zonder een stabiele interne organisatie lijkt het daarom extra moeilijk om tot goede externe samenwerkingsrelaties te komen.

Samenwerking tussen voortgezet onderwijs en welzijnswerk

Vanuit welzijnsinstellingen wordt er meer samengewerkt met het basis- dan met het voortgezet onderwijs. Toch hebben organisatienetwerken van welzijnsinstellingen en scholen voor voortgezet onderwijs hier en daar geleid tot veelbelovende gezamenlijke activiteiten. Ook hier gaat het in de meeste gevallen om samenwerking met de bibliotheek.

Positieve ervaringen met de samenwerkingen tussen VO-scholen en welzijnsinstellingen zijn in bijna alle genoemde gevallen direct gekoppeld aan activiteiten die op uitvoerend niveau zijn geformuleerd. Voor de samenwerking tussen scholen voor voortgezet onderwijs en *bibliotheken* geldt in zekere zin hetzelfde als voor de samenwerking bibliotheek-basisscholen. De psychologische afstand tussen school en bibliotheek is minder groot dan die tussen scholen en andere welzijnsinstellingen. Bij andere activiteiten liggen zaken al gauw veel gevoeliger. Zo zijn huiswerkprojecten nogal eens bron van conflicten. Dat geldt met name voor situaties waarbij vanuit het welzijnswerk lokatie en begeleiding worden geboden. Het welzijnswerk levert op die manier een bijdrage aan de instructie van leerlingen. Een competentiestrijd is in die gevallen niet uitzonderlijk.

Spijbelopvangprojecten en 'werkgroepen uitvoering leerplicht' zijn ook activiteiten waarbij voortgezet onderwijs en welzijnswerk elkaar kunnen ontmoeten. Voor veel scholen voor voortgezet onderwijs is spijbelen echter een gevoelig punt. Verzuimproblematiek wordt liever binnenskamers gehouden.

Onderwijsvoorrangsgebieden als organisatienetwerken

Vanuit het perspectief van organisatienetwerken zijn de beschreven processen die zich binnen onderwijsvoorrangsgebieden afspelen goed te interpreteren. De voor organisatienetwerken zo kenmerkende *ambiguïteit* tussen het belang van de verschillende organisaties en het belang van het netwerk is ook in de onderwijsvoorrangsgebieden goed te herkennen. Vooral tussen scholen die een opvolgende afhankelijkheidsrelatie hebben met elkaar (basisscholen en scholen voor voortgezet onderwijs; basisscholen en scholen voor speciaal onderwijs) speelt de *bescherming van de autonomie* een grote rol. Dit speelt bijvoorbeeld een rol bij de samenwerkingsproblemen in het kader van de aansluitingsproblematiek tussen basisscholen en scholen voor voortgezet onderwijs. Deze problemen zijn ook te herkennen bij samenwerking met welzijnsinstellingen. Door scholen worden deze instellingen niet altijd in hun autoriteit geaccepteerd. De machtspositie van onderwijsinstellingen is in veel gebieden sterker dan die van het welzijnswerk. De autoriteitskwestie is hier debet aan. Scholen worden in hun deskundigheid op voorhand geaccepteerd. Het welzijnswerk moet zich meer waarmaken.

Behoefte aan samenwerking blijkt er vooral te bestaan bij basisscholen onderling. De scholen hebben er belang bij organisatorische zaken samen af te stemmen. Bovendien zijn er op beleidsmatig niveau vaak minder grote verschillen.

Wanneer er *verschuiving van afhankelijkheid* optreedt ontstaan er gemakkelijk spanningen. Welzijnsinstellingen die huiswerkbegeleiding en spijbelopvangmogelijkheden bieden komen op het domein van vooral scholen voor voortgezet onderwijs. Het eigen domein wordt dan vaak beschermd, de autonomie staat centraal.

5 Besluit: voorwaarden voor goed functioneren

Voor deelname aan een onderwijsvoorrangsgebied ontvangen scholen en welzijnsinstellingen extra formatie. De inhoudelijke verantwoordelijkheid van de inzet van die formatie blijft bij het bevoegd gezag, coördinatoren van het gebied hebben formeel geen zeggenschap over de inzet van de formatie. Deelname aan een onderwijsvoorrangsgebied kan dus makkelijk voordelen opleveren (extra formatie) zonder dat daar per definitie veel autonomie voor hoeft te worden ingeleverd. Scholen en welzijnsinstellingen kunnen in zekere zin de samenwerking met andere participanten tot een minimum beperken. De coördinatoren van de gebieden hebben geen sanctionerende bevoegdheid naar scholen die hun verplichtingen niet nakomen. Een bestuur dat niet OVB-minded is blijkt de positie van de coördinator nagenoeg uit te hollen.

Een zekere inhoudelijke overeenstemming blijkt vooral aanwezig bij basisscholen onderling en basisscholen in samenwerking met welzijnsinstellingen. In sommige gevallen sluiten de werkwijzen goed op elkaar aan, herkent men elkaars problematiek en weegt het belang van de samenwerking ruimschoots op tegen de ingeleverde autonomie. Dat is niet altijd het geval. Bijvoorbeeld wanneer de verzuiling diep

verankerd is in de structuur van het betrokken gebied of wanneer scholen te maken hebben met een afnemend leerlingenaantal.

Sociale vernieuwing moet er toe leiden dat middelen voor achterstandsbestrijding doeltreffender en efficiënter worden ingezet. Afstemming van beleid van verschillende instellingen op lokaal niveau moet hier een waarborg voor zijn. Uit de ervaringen binnen onderwijsvoorrangsgebieden is gebleken dat effectieve samenwerking heel wat voeten in de aarde heeft. Instellingen scharen zich graag achter brede, vage doelstellingen als daar extra faciliteiten tegenover staan. Wanneer de doelstellingen echter in operationele termen moeten worden vertaald blijken vage doelstellingen nogal eens binnen uiteenlopende kaders van de verschillende instellingen te worden ingevuld. Doelstellingen moeten in operationele termen worden geformuleerd zodat ze binnen de mogelijkheden van afzonderlijke instellingen te realiseren zijn. Het formuleren van streefcijfers kan hierbij een hulpmiddel zijn. Het operationaliseren van doelstellingen blijkt echter vaak met problemen gepaard te gaan. De oorzaak hiervan is het feit dat er nog steeds veel onduidelijkheid is over hoe de algemene doelstellingen van het OVB op uitvoerend niveau gerealiseerd moeten worden. Als niet duidelijk is hoe onderwijsachterstanden bestreden moeten worden kan de decentralisatie van het beleid nog zo ver zijn doorgevoerd en kan er nog zo intensief worden samengewerkt, de schoolprestaties van de kinderen waar het om gaat zullen er niet beter van worden.

Literatuur

Geest, W. van (1986). *Organisatienetwerken van scholen. Handboek voor schoolmanagement en organisatie*, 4410-1 – 4410-17.

Godfroij, A. (1981). *Netwerken van organisaties*. 's-Gravenhage: VUGA.

Ministerie van Onderwijs en Wetenschappen (1985). *Wetsvoorstel OVB*. Tweede Kamer 18755 nrs. 1-3.

Ministerie van Onderwijs en Wetenschappen (1985). *Onderwijsvoorrangsplan*, Tweede Kamer 18828.

Ministerie van Onderwijs en Wetenschappen (1992). *Ceders in de tuin*. 's-Gravenhage: Sdu.

Nota Sociale Vernieuwing, Tweede Kamer, vergaderjaar 1989-1990, 21455, nrs. 3 en 4. Den Haag: Sdu, 1990.

Petersen, B., & Keijzer, M. de (1992). *Het speciaal onderwijs en het onderwijsvoorrangsbeleid*. Nijmegen: ITS.

Research voor Beleid (1987a). *Onderwijs en welzijn in onderwijsvoorrangsgebieden*. Leiden: RvB.

Research voor Beleid (1987b). *Onderwijsvoorrangsgebieden, organisatie en management*. Leiden: RvB.

Smets, P., & Kessel, N. van (1981). *Onderwijsstimuleringsprojecten: formele projectorganisatie, faciliteiten en kerncijfers*. Nijmegen: ITS.

Verhaak, C., & Brandsma, H. (1991). *Onderwijsvoorrangsgebieden op weg*. Nijmegen: ITS.

Verhaak, C., & Kristensen, D. (1992). *Projecten voor jonge allochtone kinderen*. Rijswijk: Ministerie van WVC.

Wassenberg, A. (red.) (1980). *Netwerken: organisatie en strategie*. Meppel: Boom.

Manuscript aanvaard 11-1-1993

Auteur

C. Verhaak (1959) is als onderzoeker verbonden aan het Instituut voor Toegepaste Sociale wetenschappen te Nijmegen.

Adres: Instituut voor Toegepaste Sociale wetenschappen, Postbus 9048, 6500 KJ Nijmegen,

Abstract

Cooperation between schools and social welfare-agencies to enhance educational opportunities.

C. Verhaak. Pedagogische Studiën, 1993, 70, 263-273.

In the Netherlands schools and social welfare-agencies located in areas which are characterised by a cumulation of disadvantages, may qualify to get special resources to enhance the pupils' educational opportunities. Areas in which such a cooperation is started are called: areas of educational priority. The main aim of these areas is to enhance the pupils' educational opportunities by collaboration between

schools and social welfare-agencies. Areas of educational priority should create effective programs against educational disadvantages in close harmony with local policy. In this article we report the results of a research into the state of affairs in the areas of educational priority. It appears from this study that cooperation between schools and social welfare-agencies is hard to get organised. That does not mean that within areas of educational priority a wide range of activities is being carried out. However, the aims of the activities are only vaguely described. A formulation of main aims in operational terms is mostly lacking. This is partly the result of a lack of knowledge of how to enhance educational opportunities at local level.

Visitaties lerarenopleidingen nader bekeken

F.K. Kieviet

Samenvatting

Kort na elkaar zijn drie visitatierapporten verschenen die betrekking hadden op een bepaald type lerarenopleiding. Uit een nadere beschouwing van de bevindingen van de drie visitatiecommissies blijkt dat deze bevindingen onderling sterk verschillen wat de aard ervan betreft. Deze verschillen hebben vooral betrekking op de mate van aandacht die de respectievelijke commissies hebben geschonken aan het gegeven dat het bij een lerarenopleiding om een specifieke beroepsopleiding gaat, en de mate van aandacht die men heeft gegeven aan met name onderwijskundige c.q. opleidingsdidactische aspecten van deze beroepsopleiding. Aangenomen wordt dat deze verschillen verband houden met de taakopvatting en het kader waarvan de onderscheiden commissies zijn uitgegaan. Meer onderlinge afstemming van visitaties bij lerarenopleidingen wordt wenselijk geacht, respectievelijk het maken van afspraken over de te volgen aanpak.

Inleiding

Eind 1992 respectievelijk begin 1993 verschenen kort na elkaar drie rapporten van visitatiecommissies die zich bezig hadden gehouden met een bepaald type lerarenopleiding. Deze drie rapporten waren:

- Onderwijsvisitatie universitaire lerarenopleiding (VSNU, 1992);
- Bekwaamheid aan de basis. Eindrapport van de visitatiecommissie lerarenopleiding basisonderwijs (3 delen; HBO-Raad, 1993a);
- Eindrapport van de visitatiecommissie deeltijdse eerstegraads lerarenopleiding voor de algemene vakken (HBO-Raad, 1993b).

Deze rapporten zijn er drie uit een reeks van dergelijke verslagen van visitaties van opleidingen en studies in het Hoger Beroepsonderwijs (HBO) en het Wetenschappelijk Onderwijs (WO). Deze visitaties maken deel uit van een systeem van interne en externe kwaliteitszorg in het HBO en WO.

In deze bijdrage zal eerst enige toelichting gegeven worden op het vigerende systeem van visitaties. Daarbij wordt ingegaan op visitatie als middel voor externe kwaliteitszorg (par. 1) en wordt de visitatieprocedure kort uiteengezet (par. 2). Vervolgens worden de bevindingen van de drie visitatiecommissies op hoofdpunten weergegeven (par. 3). De aard van de bevindingen van de drie visitatiecommissies blijkt nogal uiteen te lopen. Dit hangt samen met de taakopvatting en het kader waarvan de respectievelijke commissies zijn uitgegaan. Deze taakopvatting en dit kader komen aan de orde in par. 4. Ten slotte wordt nader ingegaan op de gesignaleerde verschillen en de vraag besproken of meer onderlinge afstemming van visitaties gewenst is (par. 5).

1 Externe kwaliteitszorg door middel van visitaties

Het HBO en het WO hebben zich verplicht om naast interne kwaliteitszorg ook een systeem van externe kwaliteitszorg te realiseren door middel van visitatiecommissies. Deze verplichting die men is aangegaan, houdt verband met het streven van de overheid naar deregulering en het verlangen van de instellingen om hun autonomie te vergroten. Het begrip kwaliteit staat met name in het centrum van de belangstelling sinds het verschijnen van de nota Hoger Onderwijs: Autonomie en Kwaliteit, in de wandel de HOAK-nota genoemd, in 1985 (Ministerie van Onderwijs en Wetenschappen, 1985). Autonomie van de instellingen dient volgens de filosofie van de nota gepaard te gaan met het nemen van verantwoordelijkheid voor de geleverde kwaliteit en de wijze waarop over de kwaliteit gewaakt wordt.

De HBO-Raad respectievelijk de VSNU (Vereniging van Samenwerkende Nederlandse Universiteiten) zijn volgens dezelfde filosofie

zelf verantwoordelijk voor de organisatie en uitvoering van de visitaties. Zij geven richtlijnen voor de visitaties en stellen per opleiding of studierichting de samenstelling van de visitatiecommissie vast, waarin met name externe deskundigen zitting hebben. Deze commissies bezoeken alle instellingen die een opleiding of studie in huis hebben waarop de visitatie gericht is.

Zoals aangegeven vormt dit systeem van externe kwaliteitszorg het complement van dat van interne kwaliteitszorg. Deze systemen van interne kwaliteitszorg van de instellingen zijn nog volop in ontwikkeling. De term 'zorg' die in dit verband gebruikt wordt, geeft aan dat het niet slechts gaat om vaststelling en beoordeling van de kwaliteit, - wat daaronder dan ook verstaan wordt -, maar dat verbetering van de kwaliteit het uiteindelijk doel is. Een belangrijke waarborg voor de kwaliteit van een opleiding is juist een goed functionerend systeem van interne kwaliteitszorg. De interne kwaliteitszorg wordt dan ook opgevat als een continu proces, dat periodiek wordt aangevuld met externe visitaties. Doordat de rapportage van externe visitaties openbaar is, bestaat overigens wel het gevaar dat de aandacht extern (bijvoorbeeld in de pers) en intern binnen de instelling eenzijdig uitgaat naar de visitaties en de daaruit resulterende conclusies. Dit gevaar wordt nog versterkt doordat de instellingen voorafgaande aan het bezoek van de visitatiecommissie een zogenaamde zelfstudie dienen uit te voeren. Een dergelijke zelfstudie, die moet uitmonden in een rapport ten behoeve van de visitatiecommissie, dient gericht te zijn op een analyse van de sterke en zwakke punten van de eigen opleiding. Voor deze zelfstudie zijn door de HBO-Raad respectievelijk de VSNU richtlijnen opgesteld. Niet denkbeeldig is dat instellingen hun 'kwaliteitszorg' beperken tot de uitvoering van deze periodieke zelfstudie en vervolgens het ondergaan van de visitatie. Uiteraard hebben HBO-Raad en VSNU dit gevaar zelf ook onderkend, veel belang hechten zij dan ook aan wat de instellingen in aansluiting op de visitatie met de aanbevelingen doen, als ook met de uitkomsten van hun eigen sterkte-zwakte-analyse. Vroeijenstijn (1993), die zich bij de VSNU met de visitaties bezig houdt, erkent echter dat dit vanuit de VSNU niet altijd te volgen is.

2 De visitatieprocedure

Hierboven werd reeds aangegeven dat de instellingen voorafgaande aan de visitatie een zelfstudie dienen uit te voeren waarvan zij het rapport aan de visitatiecommissie ter beschikking stellen. Een dergelijke zelfstudie moet een realistisch beeld geven van de onderwijssituatie in een bepaalde instelling; de visitatiecommissie zal ter plaatse proberen vast te stellen of dat inderdaad het geval is. De visitatiecommissie kan aan de hand van het rapport van de zelfstudie een eerste indruk krijgen en punten inventariseren waaraan zij bij haar bezoek bijzondere aandacht wil geven.

De visitatiecommissies voeren hun werk uit binnen het kader dat door de HBO-Raad respectievelijk de VSNU wordt aangereikt. Zij hebben daarbinnen een grote mate van vrijheid om hun eigen taakopvatting verder in te vullen en hun werkwijze nader te ontwikkelen. In het geval van de drie visitatiecommissies wier rapport hier ter bespreking staat, leidt dit tot interessante verschillen, die wij hierna nog nader zullen bespreken.

Bij het bezoek van de visitatiecommissie aan de betrokken instellingen voert de commissie gesprekken met (representanten van) diverse betrokkenen, waaronder ook studenten. Ook neemt zij nota van bepaalde documenten die zij ter inzage heeft gevraagd, en bekijkt zij faciliteiten zoals bibliotheek, practicavoorzieningen, audiovisuele uitrusting, computerfaciliteiten.

De verzamelde gegevens worden door de visitatiecommissie verwerkt en geanalyseerd, waarna rapportage volgt in een visitatierapport. Daarin wordt gerapporteerd over de afzonderlijke instellingen, maar ook worden conclusies getrokken over de kwaliteit van de betrokken opleiding of studie als geheel. Tevens worden aanbevelingen gedaan, zowel ten behoeve van afzonderlijke instellingen als gericht op de opleiding of studie landelijk gezien.

3 Enige bevindingen van de commissies

3.1 De Pedagogische Academie

In de dagbladpers hebben vooral de conclusies over de pedagogische academie voor het basisonderwijs (PABO) de aandacht getrokken. De gebezigde krantekoppen lieten de indruk ontstaan dat de visitatiecommissie de PABO's tot de grond toe had afgebroken. Vanuit de PABO-wereld is tegen dit negatieve beeld fel gereageerd, waarbij ook beschuldigend naar de visitatiecommissie is gewezen.

Wie het eindrapport van de visitatiecommissie lerarenopleiding basisonderwijs zorgvuldig leest, zal kunnen vaststellen dat de commissie zich genuanceerd uitdrukt, en bovendien duidelijk maakt dat de PABO-opleiding in feite nog in een ontwikkelingsfase verkeert. Zij stelt dat juist vanwege de ontwikkelingen die in de PABO-wereld gaande zijn, een definitief kwaliteitsoordeel niet goed te geven is en voorzichtigheid bij het geven van een kwaliteitsoordeel noodzakelijk is. "De visitatie moet dan ook worden gezien als een momentopname, waarbij de commissie, hoewel zij op tal van punten minder gunstig over de geleverde kwaliteit oordeelt, zeker ook ontwikkelingen in een positieve richting signaleert" (p. 10). Overigens, zo voegen wij hier aan toe, is een definitief oordeel ook onder andere omstandigheden niet op zijn plaats, wanneer men uitgaat van de eerder aangegeven gedachte dat een visitatie een bijdrage zou moeten leveren aan verbetering van de kwaliteit, het oordeel over de kwaliteit is dan allereerst formatief van aard en niet summatief.

Elders (Kieviet, 1993a) hebben wij ook zelf naar aanleiding van het visitatierapport aangegeven aan welke veranderingen de opleidingen voor het basisonderwijs onderhevig zijn geweest. In zijn proefschrift uit 1987 heeft Coonen geschetst hoe de zaken in de PABO's mede als gevolg van deze veranderingen er voor stonden; de situatie die hij aantrof karakteriseerde hij als 'een problematische vernieuwingscontext'. Uit onderzoek van vernieuwingsprocessen is genoegzaam bekend dat ingrijpende innovaties zich over een reeks van jaren uitstrekken, en in een problematische vernieuwingscontext ligt dit nog ongunstiger.

Met inachtneming van voorgaande opmerkingen is het beeld dat de visitatiecommissie lerarenopleiding basisonderwijs schetst, niet rooskleurig. Het is niet eenvoudig om de Samenvatting van acht pagina's die de commissie zelf geeft, nog verder in te dikken. Als hoofdpunten komt o.i. het volgende naar voren:

- Als algemeen oordeel stelt de commissie dat zij opleidingen heeft aangetroffen waarvan de kwaliteit redelijk tot goed genoemd kan worden. "In het algemeen leveren opleidingen op tal van aspecten die de commissie heeft onderscheiden, goed werk. De commissie constateert echter ook dat op een aantal kwaliteitsaspecten die zij als cruciaal beschouwt, de kwaliteit van veel opleidingen nog niet voldoende is. De commissie plaatst met name kanttekeningen bij de kwaliteit van het inhoudelijk opleidingskader en het beleidsvoerend vermogen van veel opleidingen" (p. 10).
- Bij het merendeel van de opleidingen ontbreekt volgens de commissie dat inhoudelijk opleidingskader met een uitgewerkt beroepsbeeld, met helder geoperationaliseerde eindtermen en een daarop aansluitende geformuleerde opleidingsvisie die richting geeft aan de curriculumontwikkeling en -evaluatie, alsmede aan het handelen van docenten.
- Er wordt over het algemeen redelijk tot goed aandacht besteed aan de vakinhouden en de vakdidactiek van de basisschoolvakken. Daar staat tegenover dat bij de meeste opleidingen onvoldoende aandacht wordt gegeven aan een aantal essentiële aspecten van het leraarschap in de basisschool, zoals het omgaan met individuele verschillen en meer meso-structurele aspecten van het leraarsberoep. De specifieke aandacht voor het onderwijs aan het jongere respectievelijk het oudere kind staat nog in de beginfase. De commissie signaleert dat vaak een functionele relatie tussen de aangeboden theorie en de praktijk ontbreekt.
- Op een deel van de opleidingen is een kwalitatief goed stageprogramma aanwezig. Anderzijds zijn er bij nogal wat opleidingen kanttekeningen te plaatsen bij het stagecurriculum. Een probleem is het gebrek aan kwalitatief goede stageplaatsen, dat men toeschrijft aan het ontbreken van een stagevergoeding.

- Een sterk punt vindt de commissie dat bij veel opleidingen de onderlinge relatie tussen docenten en studenten goed is. Daar staat tegenover dat de meeste opleidingen geen beargumenteerd en geëxpliciteerd beleid hebben voor de inrichting van het onderwijs. Het onderwijs maakt op veel opleidingen een vrij schoolse indruk en draagt doorgaans onvoldoende bij aan de ontwikkeling van zelfstandigheid en verantwoordelijkheid bij de student, zo merkt de commissie op.
- Ten aanzien van de instroom en uitstroom wordt gesteld dat veel opleidingen geen idee hebben van de aanvangskwalificaties van hun studenten, en dat veel van de opleidingen onvoldoende zicht hebben op de specifieke kwalificaties die studenten bereiken, in relatie tot de opleidingseindtermen. Het opzetten van een samenhangend systeem van interne kwaliteitszorg verkeert bij praktisch alle opleidingen nog in de beginfase.

3.2 De deeltijdse eerstegraads lerarenopleiding voor de algemene vakken

Ook over de deeltijdse eerstegraads lerarenopleiding voor de algemene vakken (dav-1) worden door de betrokken visitatiecommissie kritische opmerkingen gemaakt, maar deze kregen in de dagbladpers minder de aandacht. Voor deze opleiding geldt eveneens dat zij onderhevig was aan belangrijke ontwikkelingen. Zij is voortgekomen uit de vroegere M.O.-B opleiding. In de jaren tachtig was de opleiding onderworpen aan de STC-operatie (Schaalvergroting, Taakverdeling, Concentratie), en vond herprogrammering plaats. Met één uitzondering is dav-1 een vervolgopleiding, d.w.z. dat alleen studenten worden toegelaten die reeds een tweedegraads lerarenopleiding hebben afgesloten (of een andere daarmee vergelijkbare opleiding hebben genoten). In verband daarmee gaat men ervan uit dat de zogenaamde beroepsvoorbereiding al voor een groot deel heeft plaatsgevonden. Niettemin is dav-1 een beroepsopleiding voor het leraarschap, en wel op eerstegraads niveau.

De hoofdpunten van het oordeel van de visitatiecommissie zijn de volgende:

- De commissie is in het algemeen van oordeel dat dav-1, ondanks de grote inzet van alle betrokkenen en de tevredenheid van het merendeel van de studenten, nauwelijks voldoet aan de eisen die zij in haar referentiekader had geformuleerd. Deze eisen van het referentiekader hadden betrekking op de doelstellingen, de kwaliteit van de vakopleiding, de kwaliteit van de beroepsvoorbereiding, de verhouding tussen vakopleiding en beroepsvoorbereiding, het beleid en de organisatie, en de kwaliteitszorg.
- De commissie stelt dat de grote meerderheid van de opleidingen niet beschikt over opleidingsdoelen die gerelateerd zijn aan een expliciet en richtinggevend beeld van de leraar. De voorbereiding op belangrijke aspecten van het leraarschap vindt bij de meeste opleidingen hoofdzakelijk impliciet en niet-systematisch plaats. Alle opleidingen geven aan het vakgerichte deel van de opleiding de hoogste prioriteit. De commissie vindt de kwaliteit daarvan in het algemeen redelijk. Anderzijds constateert de commissie dat dit deel weinig interne samenhang vertoont. Slechts een klein deel van de opleidingen is begonnen met een concrete afstemming van sommige programmadelen van de vakopleiding op de beroepswerkzaamheid van leraren. De werkvormen zijn bij de meeste opleidingen niet afgestemd op volwassenen met beroepservaring als leraar, en evenmin op ontwikkelingen in de schoolpraktijk en de taken van leraren daarin.
- Bij het merendeel van de opleidingen is slechts voor een deel een streven zichtbaar naar afstemming van de beroepsvoorbereiding op wat leraren in het eerstegraads gebied nodig hebben, aldus de commissie. Bij het merendeel van de opleidingen ontbreekt een systematisch overzicht van de kenmerkende leraarstaken in het eerstegraads gebied. Dat geldt ook voor het element van reflectie op de eigen stijl die de studenten in hun eigen praktijk ontwikkeld hebben. Ten gevolge van het gebrek aan een overzicht van beroepsvereisten zijn, aldus de commissie, de meeste opleidingen niet in staat vast te stellen welke inhoud en omvang de directe beroepsvoorbereiding moet hebben.
- De commissie constateert dat de dav-1 in de organisatie en het beleid van de HBO-instellingen slechts een marginale plaats inneemt.

- De commissie acht de resultaten van de opleidingen in kwantitatief opzicht teleurstellend. Zo staakt ruim 40% van de studenten de studie voortijdig. De interne kwaliteitszorg staat bij dav-1 nog in de kinderschoenen.

De visitatiecommissie verbindt aan haar bevindingen de uitspraak dat de deeltijdse eerstegraads lerarenopleidingen voor de algemene vakken niet in een positie verkeren en evenmin over de middelen beschikken om aan de eisen van het eerder genoemde referentiekader te kunnen voldoen, vooral als gevolg van hun kleinschaligheid en historische achtergrond. Voor de verdere ontwikkeling van deze opleidingen, die de commissie wel van belang acht, is volgens de commissie een herziening van de huidige constellatie ("reconstructie") noodzakelijk.

3.3 De universitaire lerarenopleiding

Wat geldt voor de bovenbesproken opleidingen gaat eveneens op voor de universitaire lerarenopleiding (ulo): ook zij was in de achterliggende jaren onderhevig aan veranderingen èn onzekerheden. De universitaire lerarenopleiding werd een tweede-fase-opleiding in de omvang van één jaar in aansluiting op een predoctorale cursus van twee maanden. De helft van de cursusduur dient daarbij besteed te worden aan een stage (het schoolpracticum).
De bevindingen van de visitatiecommissie komen op het volgende neer:

- "De visitatiecommissie is van oordeel dat de kwaliteit van de tweede-fase universitaire lerarenopleiding globaal gesproken varieert van middelmatig tot zeer goed, zij het dat ze voor verbetering vatbaar is. De aangetroffen curricula en werkwijzen, en de betrokkenheid van docenten ondersteunen het oordeel van de Commissie. De studenten en afgestudeerden zijn tevreden over hun opleiding. Die tevredenheid wordt gedeeld door het afnemende veld. Niettemin heeft de Commissie een aantal zwakten aangetroffen" (p. 34). Deze worden door de Commissie niet nader gepreciseerd.
- De commissie signaleert een schril contrast tussen de kleinschaligheid die de universitaire lerarenopleidingen kenmerkt, en de complexiteit van de structuren. De kleinschaligheid zelf acht de commissie als zodanig ook bedreigend.
- De commissie constateert dat het beroepsbeeld dat sedert een daarop betrekking hebbend advies uit 1986 in de opleidingen wordt gehanteerd, in de meeste gevallen niet of nauwelijks geëvolueerd is. In een aantal gevallen functioneren binnen een opleiding verschillende modellen, overigens zonder veel evaluatie ervan, naast elkaar.
- Volgens de commissie wordt de wetenschaps- en onderzoekscomponent in de universitaire lerarenopleiding als gevolg van een aantal omstandigheden bedreigd.
- De aansluiting tussen de pre-doctorale cursus en de tweede fase opleiding laat structureel en soms ook inhoudelijk te wensen over.

4 Taakopvatting en het gebruikte kader van de visitatiecommissies

In het voorgaande hebben we de bevindingen van de drie visitatiecommissies op hoofdpunten aangegeven. Wij hebben daarbij gebruik gemaakt van de samenvattingen die de commissies zelf in hun respectievelijke rapporten gepresenteerd hebben, en zijn zo dicht mogelijk bij de oorspronkelijk tekst gebleven.

Wanneer de bevindingen die we hierboven weergaven, onderling vergeleken worden, valt een aantal verschillen op in de aard ervan. Dit houdt geen verband met de wijze waarop wij de bevindingen hebben weergegeven, dat is een te gemakkelijke verklaring. De verschillen hebben vooral betrekking op een tweetal punten:

- de mate van aandacht die gegeven wordt aan het gegeven dat het bij een lerarenopleiding om een bepaald type beroepsopleiding gaat;
- de mate van aandacht die geschonken wordt aan met name onderwijskundige of opleidingsdidactische kanten van de beroepsopleiding in verhouding tot andere aspecten.

O.i. houdt deze verschillende aandacht verband met de taakopvatting en het kader waarvan de respectievelijke commissies zijn uitgegaan. In het navolgende zullen we deze

taakopvatting en het gehanteerde kader in hoofdlijnen aangeven. Wij maken daarbij gebruik van een eerdere beschrijving daarvan (Kieviet, 1993b).

Het rapport van de visitatiecommissie lerarenopleiding basisonderwijs onderscheidt zich doordat een zeer uitvoerige verantwoording wordt gegeven van de aanpak. De commissie ziet visitatie als middel voor kwaliteitsverbetering, met name van de afzonderlijke opleidingen. Mede in verband daarmee is zij bij de beoordeling uitgegaan van proceskwaliteit, waarbij het gaat om de vraag of de voorwaarden die noodzakelijk geacht worden voor een goed 'eindprodukt', aanwezig zijn. De commissie heeft zichzelf daarbij gezien als 'consultant': uitgaande van de beschikbare 'expertknowledge' wordt geprobeerd zich in te leven in de situatie van de 'cliënt' en op basis van deze kennis een diagnose te stellen om vervolgens een aantal aanbevelingen ter verbetering van de situatie te formuleren.

De visitatiecommissie staat een dynamische kwaliteitsopvatting voor. Volgens haar is het niet mogelijk om standaarden te hanteren met een universele geldigheidswaarde. Wel is een aantal eisen te formuleren die aan iedere opleiding gesteld mogen en kunnen worden. Een deel van de operationalisatie van kwaliteit is echter afhankelijk van het type opleiding waarom het gaat, in casu een lerarenopleiding, en ook is de invulling afhankelijk van het ontwikkelingsstadium waarin de opleiding verkeert.

Het kader waar de commissie bij de beoordeling van de kwaliteit van is uitgegaan, bevat vier componenten:

- vier kwaliteitsperspectieven (een juridisch-taakstellend, een opleidingsdidactisch, een bedrijfseconomisch en een organisatie-ontwikkelingsperspectief);
- negen objecten van kwaliteitsbeoordeling, te weten: het inhoudelijk opleidingskader (uitgangspunten, beroepsbeeld, eindtermen en doelen), het curriculum, het onderwijsleerproces, de studievoortgang van studenten, beleid en organisatie, docenten en personeelsbeleid, de interne kwaliteitszorg, de onderwijsresultaten en de nascholing;
- één of meer evaluatievragen per kwaliteitsobject uitgaande van de vier kwaliteitsperspectieven;
- een specificatie van de evaluatievragen in een aantal ijkpunten. Deze ijkpunten waren voor de commissie de concrete aandachtspunten voor de beantwoording van de evaluatievragen; zij vormden een operationalisatie van de kwaliteit zoals die door de commissie is opgevat. Een voorbeeld van zo'n ijkpunt voor het object 'het inhoudelijk opleidingskader' was b.v.: Er is een inhoudelijke relatie tussen beroepsbeeld en de eindtermen.

De commissie heeft volgens haarzelf bij het formuleren van de ijkpunten terdege rekening gehouden met de beschikbare kennis en inzichten over de eisen die aan kwalitatief goed onderwijs mogen worden gesteld. Het is denkbaar dat men het op bepaalde punten niet eens is met de commissie, het voordeel van haar aanpak is wel dat in ieder geval discussie over haar operationalisatie van kwaliteit mogelijk is.

De visitatiecommissie deeltijdse eerstegraads lerarenopleiding voor de algemene vakken heeft zichzelf bij voorkeur willen beschouwen als een kritische gesprekspartner voor de opleidingen, met als doel de kwaliteit van de opleidingen en de zorg ervoor te bevorderen. De commissie heeft in eerste instantie een lijst opgesteld van aandachtspunten die de agenda vormden voor de bestudering van de zelfevaluatierapporten en voor de gesprekken bij de visitatiebezoeken. In aansluiting hierop heeft de commissie een referentiekader ontwikkeld voor de beoordeling van de kwaliteit van de opleidingen. In het bovenstaande werden de categorieën van dit referentiekader reeds genoemd. Binnen dit referentiekader heeft de commissie algemene eisen geformuleerd waaraan de opleidingen volgens haar zouden moeten voldoen. Daarbij is het uitgangspunt geweest dat dav-1 een opleiding is die voorbereidt op het beroep van leraar in het eerstegraads gebied. Voor de vormgeving van een dergelijke beroepsopleiding is het noodzakelijk inzicht te hebben in de eisen die het beroep stelt en de kwalificaties waarover de beroepsbeoefenaar dient te beschikken.

Ook de visitatiecommissie voor de universitaire lerarenopleiding gaat ervanuit dat de opleiding gezien moet worden als beroepsopleiding: het gaat om de opleiding van een hoogwaardig type van leraar, 'professionals'. Als inhoud van het beroepsprofiel worden door

de commissie vier zeer algemeen geformuleerde punten aangegeven, die min of meer overeenkomen met die uit de nota van de ARCULO (Academische Raad, Commissie Universitaire Lerarenopleiding) waarvan de opleidingen uitgaan. De eisen die de visitatiecommissie formuleert waaraan de initiële opleiding moet voldoen, zijn:

- "Het curriculum moet erop gericht zijn de kennis en vaardigheden ten behoeve van het geschetste beroepsprofiel te ontwikkelen.
- De docenten van de lerarenopleiding moeten hun onderwijs stoelen op een gedegen en brede kennis van de betrokken wetenschapsgebieden en zelf aan het onderzoek participeren" (p. 5).

Met het bovenstaande zijn de taakopvatting en het gebruikte kader van de visitatiecommissie universitaire lerarenopleiding aangegeven.

5 Naar meer afstemming

Eerder hebben wij gesteld dat de verschillen in de aard van de bevindingen enerzijds betrekking hebben op de mate waarin men aandacht heeft geschonken aan het gegeven dat het bij een lerarenopleiding gaat om een specifieke beroepsopleiding, anderzijds op de mate van aandacht die men heeft gegeven aan met name onderwijskundige of opleidingsdidactische aspecten van deze beroepsopleiding. Deze verschillen zouden verband houden met de taakopvatting en het kader waar de onderscheiden commissies vanuit zijn gegaan.

Wanneer men de taakopvatting en het kader vanwaaruit door de commissies gewerkt is, onderling vergelijkt, tekenen zich twee uitersten af. Het meest expliciet in haar taakopvatting is de visitatiecommissie lerarenopleiding basisonderwijs. Ook geeft zij een sterk geleed en uitgewerkt kader aan voor de kwaliteitsbeoordeling van de PABO als lerarenopleiding. Aan het andere uiterste is de visitatiecommissie universitaire lerarenopleiding te lokaliseren: de taakopvatting wordt niet aangegeven, terwijl het kader uiterst globaal blijft.

Gegeven de aangegeven expliciete taakopvatting en het gehanteerde uitgewerkte kader van de PABO-visitatiecommissie kan men een uitgebreid verslag van bevindingen verwachten waarin met name vele verschillende onderwijskundige of opleidingsdidactische aspecten in rechtstreekse relatie met de PABO als beroepsopleiding aan de orde komen. Wanneer men de bevindingen van de ulo-visitatiecommissie nagaat, valt op dat deze commissie nauwelijks opmerkingen maakt die betrekking hebben op de inhoud en de vormgeving van de ulo als beroepsopleiding. De onderwerpen die de commissie in haar verslaggeving aansnijdt hebben voor het merendeel betrekking op randvoorwaardelijke aangelegenheden (waarmee niet gezegd wil zijn dat die niet belangrijk zouden zijn). De algemene uitspraak die de ulo-visitatiecommissie over het onderwijs doet is in feite niet gebaseerd op informatie over de primaire processen van het onderwijs zelf of over de resultaten van dit onderwijs (zie ook Hoeben, 1993). De visitatie van de universitaire lerarenopleiding voldoet ons inziens niet aan de eisen die onderwijskundig gezien aan een adequate doorlichting van een lerarenopleiding gesteld zouden mogen worden.

In het begin van deze uiteenzetting hebben wij aangegeven dat visitatiecommissies binnen het door de HBO-Raad of VSNU aangereikte kader een grote mate van vrijheid hebben om hun eigen taakopvatting verder in te vullen en hun werkwijze nader te ontwikkelen. Blijkbaar kan deze relatief grote mate van vrijheid er toe leiden dat in het geval van lerarenopleidingen sterk verschillende taakopvattingen en kaders gehanteerd worden, waardoor het kwaliteitsoordeel op uiteenlopende zaken betrekking heeft, en verschillende kwaliteitscriteria worden aangelegd. Dit betekent dat onderscheiden typen van lerarenopleidingen op sterk uiteenlopende wijzen zijn beoordeeld, zonder dat dit overigens verband houdt met de specifieke context van die typen van opleiding.

Wij achten dit een ongewenste situatie. In de eerste plaats is een opleiding waarover uitvoerig gerapporteerd wordt, onder meer wat inhoud en vormgeving van het programma betreft, voor de buitenwereld, met name de media, meer kwetsbaar. Maar wat belangrijker is: er kan een gemaskeerd beeld ontstaan van hoe de opleiding als beroepsopleiding werkelijk functioneert. Bovendien worden de instellingen tekort gedaan wanneer in de evaluatie door een visitatiecommissie belangrijke elementen buiten beschouwing worden gelaten: zij hebben dan ook niet de mogelijkheid te pro-

aanwijzingen om eventuele tekorten in de opleiding op te heffen.

Wij zijn van mening dat er in de toekomst bij herhaling van visitaties van lerarenopleidingen afspraken gemaakt dienen te worden om een zekere afstemming te verkrijgen ten aanzien van de te volgen aanpak. Met opzet spreken wij van afspraken, omdat een algemeen aanvaard model voor de evaluatie van lerarenopleidingen in de literatuur niet beschikbaar is (Kieviet, 1989). Wel worden veelal drie terreinen aangewezen waarop het kwaliteitsbegrip en daarmee ook de evaluatie betrekking zouden kunnen c.q. moeten hebben (Tillema, 1991), te weten de kwaliteit van de instroom van studenten, de kwaliteit van de uitstroom, en de kwaliteit van de opleiding. Hoeben (1993) gaat zelfs zover dat hij de kwaliteit van de uitstroom als het ultieme criterium voor de kwaliteit van de opleiding aanwijst. De visitatiecommissie lerarenopleiding basisonderwijs heeft, zoals gezegd, uitdrukkelijk gekozen voor de beoordeling van proceskwaliteit, niet omdat zij produktkwaliteit niet belangrijk vindt, maar omdat hier praktische problemen opdoemen: er zijn geen gecentraliseerde examens of gestandaardiseerde instrumenten (meer) waarmee hogescholen op overeenkomstige wijze de kwalificaties van afgestudeerden beoordelen. Dat is één praktisch probleem, en zo zijn er meerdere. Vandaar dat zij heeft gekeken naar de kwaliteit van de processen waarmee het 'eindprodukt' tot stand komt.

Bij laatstgenoemd punt zouden we wel willen aantekenen dat een belangrijk aandachtspunt bij visitaties dan wel dient te zijn de wijze waarop bij voortduring door de opleiding wordt nagegaan in hoeverre de student de gestelde doelen heeft bereikt. Dat veronderstelt dat de eindtermen van de opleiding vergaand geoperationaliseerd zijn en dat een verscheidenheid van testsituaties beschikbaar is om de vooruitgang van studenten te toetsen. Als we zeggen: een verscheidenheid van testsituaties bedoelen we daarmee aan te geven dat paper-and-pencil tests niet volstaan, maar dat gebruik gemaakt wordt van ecologisch zo valide mogelijke instrumenten, waarbij gebruik gemaakt wordt van bijvoorbeeld simulatie en videotapes (Kieviet, 1989). Het spreekt vanzelf dat bij de visitatie ook inzicht geboden dient te worden in de resultaten van de toetsingen, alsmede het uiteindelijk rendement.

Ten aanzien van het vaststellen van de kwaliteit van de instroom doet zich eveneens een groot aantal problemen voor. Zo is vooralsnog onduidelijk welke gegevens van beginnende studenten relevant zijn buiten gegevens over mogelijke achterstanden in vakkennis (zie de beschouwing over de mogelijkheid van selectie in Kieviet, 1989).

Ons pleidooi voor meer afstemming van visitaties bij verschillende typen van lerarenopleidingen op elkaar c.q. voor afspraken over de te volgen aanpak heeft niet de bedoeling te sturen in een richting waarbij visitaties gebruikt worden voor summatieve evaluatie, d.w.z. het uitspreken van een absoluut of vergelijkend oordeel over de kwaliteit van typen van opleidingen of individuele opleidingen. De aard van visitaties en het overwegend kwalitatieve materiaal dat daarbij aan de orde is, lenen zich o.i. slechts voor formatieve evaluatie, gericht op de bevordering van de kwaliteit van de opleiding. Voorstellen in de richting van het uitdelen van cijfers voor kwaliteit of het opstellen van een rangorde van instellingen wijzen wij af.

Literatuur

Coonen, H.W.A.M. (1987). *De opleiding van leraren basisonderwijs*. Dissertatie Rijksuniversiteit Leiden.

HBO-Raad. (1993a). *Bekwaamheid aan de basis*. Eindrapport van de visitatiecommissie lerarenopleiding basisonderwijs (3 delen). 's-Gravenhage: HBO-Raad.

HBO-Raad. (1993b). *Eindrapport van de visitatiecommissie deeltijdse eerstegraads lerarenopleiding voor de algemene vakken*. 's-Gravenhage: HBO-Raad.

Hoeben, W. (1993). De kwaliteit van universitaire lerarenopleidingen en de professionalisering van het leraarsberoep. *VELON Tijdschrift voor Lerarenopleiders, 14*(3), 10-17.

Kieviet, F.K. (1989). Kwaliteitsbewaking, evaluatie en selectie in de opleiding van onderwijsgevenden. *Pedagogische Studiën, 66*, 480-489.

Kieviet, F.K. (1993a). De PABO doorgelicht. *Jeugd en Samenleving* (in druk).

Kieviet, F.K. (1993b). Visitaties. *Vernieuwing, Tijdschrift voor Onderwijs en Opvoeding, 52*(5), 20-21.

Ministerie van Onderwijs en Wetenschappen. (1985). *Hoger Onderwijs: Autonomie en Kwaliteit.* 's-Gravenhage: Staatsuitgeverij.

Tillema, H.H. (1991). Kwaliteit van de lerarenopleiding: Een analytische evaluatie. *Tijdschrift voor Hoger Onderwijs, 9*, 19-30.

Vroeijenstijn, A.I. (1993). Zorg voor onderwijskundige kwaliteit in het hoger onderwijs. In L.P.T. Raijmakers (Red.), *Zorg voor pedagogische kwaliteit (pp. 55-67).* Assen: Dekker & Van de Vegt.

VSNU. (1992). *Onderwijsvisitatie universitaire lerarenopleiding.* Utrecht: VSNU.

Manuscript aanvaard 2-7-1993

Auteur

F. K. Kieviet is hoogleraar aan de Rijksuniversiteit Leiden

Adres: Rijksuniversiteit Leiden, Vakgroep Onderwijsstudies, Postbus 9555, 2300 RB Leiden

Abstract

Visitation reports on teacher education institutes reviewed

F. K. Kieviet. Pedagogische Studiën, 1993, 70, 274-282.

The visitation reports of three different visitation committees on three kinds of teacher education institutes have recently been published. The findings of these three committees differ strongly in nature. These differences concern the attention paid to the professional character of teacher education and the educational aspects of teacher training. The divergences seem to result from differences in the committees' views of their task and in the perspectives from which they worked. Better coordination of approaches to visitation is recommended.

Rendementsbekostiging in de bovenbouw van HAVO en VWO; overwegingen en reacties

H. W. C. H. van Amelsvoort, R. J. Bosker en T. F. Brandsma

Samenvatting

De overheid overweegt de invoering van een vorm van rendementsbekostiging voor de bovenbouw van HAVO- en VWO-scholen als middel om scholen te stimuleren hun rendement te verbeteren. Verschillen in rendement tussen scholen kunnen voor het grootste deel worden toegeschreven aan verschillen in de leerlingpopulatie. De mate van invloed die scholen via beleidsmaatregelen kunnen uitoefenen op de hoogte van het rendement mag niet overschat worden. Het is bovendien nog onduidelijk welke beleidsmaatregelen tot rendementsverbeteringen kunnen leiden. Zowel uit buitenlandse studies als uit een onder 53 HAVO- en VWO-scholen uitgevoerd onderzoek blijkt dat scholen onder invloed van bekostigingsprikkels wel geneigd zijn tot het treffen van strategische beleidsmaatregelen die gericht zijn op rendementsverbetering. Hoewel de scholen over het algemeen negatief oordelen over invoering van rendementsbekostiging, is niet aangetoond dat scholen ontwijkingsgedrag zullen vertonen dat gericht is op een vorm van budgetmaximalisatie die indruist tegen de belangen van de leerlingen.

Inleiding

Het rendement van het Nederlandse onderwijs vormt in toenemende mate een aandachtspunt in het overheidsbeleid. Ervaren knelpunten ten aanzien van de bovenbouw van het HAVO en het VWO betreffen de voortijdige uitval, de verblijfsduur, het verschijnsel van omwegen en de ontwikkeling van het rendement. Momenteel ligt het rendement van de bovenbouw van het HAVO op 74% en dat van de bovenbouw VWO op 82%. De ongediplomeerde uitval in het HAVO is de laatste jaren gestegen tot 21% en in het VWO tot 17%. Er is sprake van een discrepantie tussen de nominale studieduur en de verblijfsduur. Zo is de gemiddelde verblijfsduur in de bovenbouw van het HAVO 2,3 jaar en in het VWO 3,3 jaar. Vanuit het HAVO gingen in 1987 24% van de jongens en 22% van de meisjes naar het MBO; vanuit het VWO gingen in datzelfde jaar 22% van de jongens en 28% van de meisjes naar het HBO. In zekere zin zijn dit echter omwegen daar het MBO en het HBO ook te bereiken zijn via respectievelijk MAVO-4 en HAVO-5. Een toename van ondoelmatige leerwegen treedt eveneens op doordat steeds meer MBO- en HBO-gediplomeerden instromen in respectievelijk het HBO en het WO. Een groeiend beroep op door de overheid bekostigd onderwijs als gevolg van ondoelmatige leerwegen, staat op gespannen voet met de beperkt beschikbare overheidsmiddelen (Ministerie van Financiën, 1991).

Eén van de beleidsmaatregelen die de overheid voor de bovenbouw van het HAVO en het VWO overweegt is het invoeren van een vorm van rendementsbekostiging, waarbij de bekostiging gedeeltelijk afhankelijk wordt gesteld van het behaalde rendement. Ze beoogt hiermee enerzijds de zorg voor een betere determinatie en het terugdringen van zittenblijven en uitval te ondersteunen en anderzijds om doorstroom en gediplomeerde uitstroom te stimuleren.

De assumptie die ten grondslag ligt aan rendementsbekostiging is dat scholen in staat zijn het eigen rendement te verhogen en dat financiële prikkels hiertoe een stimulans vormen. Het is vooralsnog onduidelijk of invoering van een vorm van rendementsbekostiging scholen zal bewegen tot het voeren van een strategisch beleid dat leidt tot rendementsverbetering en vergroting van de onderwijskansen. Een potentiëel gevaar schuilt in het mogelijkerwijs optreden van strategisch of ontwijkingsgedrag. Scholen zouden maatregelen kunnen treffen om hun kwantitatief rendement ten koste van

hun kwalitatief rendement te verhogen. Hiervan is bijvoorbeeld sprake als het percentage geslaagden toeneemt of leerlingen sneller hun diploma behalen, terwijl het geheel aan kennis en vaardigheden dat de leerlingen zich hebben eigen gemaakt, afneemt. Ook zouden scholen hun rendement kunnen verbeteren door de instroomselectie te verscherpen. Invoering van een systeem van rendementsbekostiging wordt derhalve alleen als een reële optie beschouwd, als het uitblijven van dergelijk ontwijkingsgedrag voldoende kan worden gegarandeerd.

In het onderhavige artikel worden de resultaten van een onderzoek naar deze materie gerapporteerd[1] (Van Amelsvoort, Bosker, & Brandsma, 1992). Centraal in dit onderzoek staat de vraag naar het oordeel van schoolleiders over rendementsbekostiging en naar de mogelijke beleidsreacties binnen de school bij invoering van vier gespecificeerde bekostigingsscenario's. Allereerst wordt in de volgende paragrafen een conceptueel kader geschetst rond de begrippen onderwijseffectiviteit, onderwijsoutput, outputcontrole, bekostigingsprikkels en schoolbeleid. Tevens zal aandacht worden besteed aan hun onderlinge samenhang, waarbij met name de manipuleerbaarheid van het rendement zal worden toegelicht. Vervolgens worden het onderzoek en de resultaten beschreven. Daarbij wordt antwoord gegeven op de vragen: 1) Hoe oordelen schoolleiders over invoering van rendementsbekostiging? 2) Welke maatregelen in termen van strategisch gedrag of strategisch beleid zouden scholen treffen indien de beoogde rendementsbekostiging wordt ingevoerd? 3) Welke veranderingen treden op in aspecten van het huidige onderwijskundige beleid en in welke mate hangen de verschillen tussen scholen samen met het feitelijke rendement? Het artikel eindigt met een reflectie op de gevonden onderzoeksresultaten in het licht van de geschetste problematiek.

1 Theoretisch kader

1.1 Onderwijseffectiviteit en outputcontrole

Met een financiering op basis van rendement verplaatst de aandacht van de overheid zich per definitie van de input en procesbewaking naar de outputkant van het onderwijs. Outputcontrole is nauw verwant aan het denken over de kwaliteit van het onderwijs in termen van opbrengsten, effectiviteit en kosten en baten.

Twee cruciale vragen komen op bij het denken over outputcontrole: 1) Waardoor kunnen verschillen in output tussen scholen worden verklaard en 2) Wat kunnen scholen doen om hun output, i.c. hun rendement, te verbeteren? Een antwoord op deze twee vragen is van belang omdat duidelijkheid over de mate en aard van de invloed die een school op haar rendement kan hebben, een legitimatie kan vormen voor een systeem van aan outputcontrole gekoppelde bekostigingsprikkels.

De opbrengst van het onderwijseffectiviteitsonderzoek biedt een relevante kennisbasis voor het antwoord op de eerste vraag. Inputverschillen, en dan met name leerlingkenmerken als intelligentie en sociaal-economische achtergrond, verklaren verreweg het grootste deel van de verschillen in output tussen scholen. Het rendement van een school wordt met andere woorden vooral bepaald door de kenmerken van de aanwezige leerlingpopulatie, die voor veel scholen een gegeven is. Onderzoeksgegevens uit het onderwijseconomisch input-outputonderzoek verschaffen geen eenduidigheid omtrent de relatie tussen financiële input/onderwijsuitgaven en onderwijsoutput (Brempong & Gyapong, 1991; Hanushek, 1986; McCarty & Brazer, 1990; Walberg, 1992). Een mogelijke verklaring voor deze onduidelijke relatie is dat slechts een gering deel van de financiële middelen die ter beschikking worden gesteld daadwerkelijk de leerling bereikt; minder dan eenderde van het bestede geld wordt op klasniveau aan de leerlingen besteed. Overigens stelt Monk (1991) kritisch dat dit type onderzoek zich te weinig heeft gericht op de verschillen die tussen leerlingen kunnen optreden met betrekking tot de middelen die rechtstreeks aan hen ten goede komen. Tot op heden is geen duidelijke directe relatie tussen financiële middelen en onderwijsoutput aangetoond; de ene school heeft in relatie tot de middelen die ze verbruikt, hogere opbrengsten dan een andere school.

Organisatorische en onderwijsinhoudelijke schoolkenmerken blijken ook van invloed op de effectiviteit van een school. Het gaat hier om schoolspecifieke kenmerken die onder invloed van het eigen schoolbeleid gestalte krijgen. In-

dien gekeken wordt naar de sterkte van de samenhang tussen schoolkenmerken enerzijds en de output van scholen anderzijds, in termen van de variantie in leerlingprestaties, dan blijkt dat schoolkenmerken circa 11% van die variantie verklaren. Omgerekend naar tijd, impliceert dit cijfer dat een gemiddelde leerling op de 20% meest effectieve scholen qua leerprestaties een heel leerjaar voorligt op de gemiddelde leerling van de 20% minst effectieve scholen (Bosker, Guldemond, Hofman & Hofman, 1989).

Aldus kunnen we op de eerste vraag antwoorden dat vooral de kenmerken van de leerlingpopulatie, maar gedeeltelijk ook de schoolkenmerken verantwoordelijk zijn voor het gerealiseerde rendement, terwijl de invloed van financiële middelen onduidelijk is. In principe kunnen scholen hun rendement dus manipuleren, indien ze weten welke onderwijsinhoudelijke en organisatorische maatregelen daartoe genomen moeten worden. Hiermee komen we bij de tweede vraag naar wat scholen kunnen doen om hun rendement te verbeteren.

Het onderzoek naar effectieve scholen levert een aantal context- en procesfactoren op dat herhaaldelijk blijkt samen te hangen met effectiviteit. Het gaat dan om hoge verwachtingen ten aanzien van de prestaties van de leerlingen, een ordelijk klimaat gericht op leren, sterk onderwijskundig leiderschap, nadruk op basisvaardigheden, frequente evaluatie (Scheerens & Creemers, 1989), het bestaan van een gezamenlijke missie, voldoende gelegenheid tot leren en betrokkenheid van ouders en omgeving (Webster, Mendro & Almaquer, 1992).

Een problematisch aspect van dit schooleffectiviteitsonderzoek betreft het causaliteitsvraagstuk. Zowel het overwegend correlationele karakter van veel studies als een ontoereikende theorievorming ter verklaring van de invloed van bepaalde schoolkenmerken op de effectiviteit, is hier debet aan. Gezien de huidige stand van zaken in het onderwijseffectiviteitsonderzoek wordt geconcludeerd dat er weliswaar samenhang is geconstateerd tussen kenmerken op organisatie-, curriculum- en instructieniveau enerzijds en effectiviteit en rendement anderzijds, maar dat het inzicht in de verwevenheid, alsmede de relatieve bijdrage van deze kenmerken geenszins volledig is. We kunnen, met andere woorden, scholen nog geen duidelijk antwoord geven op de vraag wat ze moeten doen om hun rendement te verbeteren.

Welke consequenties heeft deze conclusie nu voor een systeem van outputcontrole en een daaruit voortvloeiende outputfinanciering? Monk (1991) concludeert dat we hier te maken hebben met een merkwaardige tegenstrijdigheid. De overheid concentreert zich op de resultaten van het onderwijs (vanuit een optimisme over de vooruitzichten om het rendement van het onderwijs te vergroten) en legt de verantwoordelijkheid voor de produktie van de prestaties bij de scholen, zonder dat duidelijkheid bestaat over de factoren die de effectiviteit bevorderen, ofwel zonder dat duidelijkheid bestaat over de produktiefunctie van de school. Daarnaast wijst Monk erop dat de overheid outputcontrole of rendementsbekostiging natuurlijk ook kan hanteren als middel om meer inzicht te krijgen in de aard van de onderwijsproduktiefunctie. In dat geval liggen er echter ook overheidstaken op het gebied van onderzoek, ontwikkeling en de verspreiding van ervaringen. Belangrijk is dan dat er een evenwicht wordt gevonden tussen de mate van deregulering en decentralisatie en de verantwoordelijkheid van de overheid voor deze taken teneinde het rendement van het onderwijs daadwerkelijk te vergroten. Het belang van dit evenwicht kan met twee extreme situaties worden toegelicht. Bij een te grote autonomie van de scholen en een te eenzijdige aandacht van de overheid voor de output, bestaat de kans dat enkele scholen succesvolle (i.c. rendementsbevorderende) combinaties ontdekken. Er is echter geen garantie dat de inzichten worden verspreid en er bestaat zelfs de kans dat de kennis wordt afgeschermd vanwege de beloning. Een te grote overheidsinmenging, waarbij de aandacht van de overheid zich ook op de proceskenmerken richt, is daarentegen denkbaar wanneer de determinanten van effectiviteit langzamerhand bekend worden; de overheid kan dan 'de beste methode' gaan voorschrijven om het rendement van het onderwijs te vergroten. De outputbenadering raakt dan achterhaald (Monk, 1991).

1.2 Bekostigingsprikkels

We hebben gezien dat er de nodige terughoudendheid dient te worden betracht bij de verwachting dat scholen in staat zijn hun rendement te verbeteren. In hoeverre geld

desalniettemin een adequate prikkel vormt tot gedrags- en/of beleidsveranderingen in het perspectief van verbetering van het rendement van de school, wordt hierna toegelicht met enkele voorbeelden van ervaringen met bekostigingsprikkels. We maken daartoe onderscheid tussen het belonen van prestaties van docenten en het belonen van prestaties van scholen.

1.2.1 Ervaringen met prestatiebeloning van docenten

Op het niveau van de docent kennen we in Nederland een systeem van uniforme betaling, waarin geen financiële beloning voor excellente prestaties en geen financiële straf voor ondermaatse prestaties is voorzien (Van der Bij, Van Noort & Vrancken, 1993). Enkele ervaringen met prestatiebeloning uit het buitenland zijn wel voorhanden. Er zijn aanwijzingen dat leerkrachten inderdaad hun gedrag veranderen onder invloed van bekostigingsprikkels. Murmane en Cohen (1986) onderzochten systemen van prestatiebeloning van docenten, waarbij de prestaties werden afgemeten aan de toegenomen leerlingprestaties of aan het oordeel van schoolleiders.

In het eerste geval doemen vragen op als "Moeten de gemiddelde toegenomen leerlingprestaties in één of in alle vakken worden bekeken?" en "Zal de leerkracht zijn tijd vooral gaan besteden aan die leerlingen die de gemiddelde score kunnen opvijzelen?". De vragen duiden op de noodzaak van het ontwikkelen van een zeer verfijnd systeem. Dit nu levert juist problemen op. Het eerste probleem heeft te maken met een gebrek aan consensus over de doelen en over wat er dus moet worden beloond. Zelfs als er over de doelen consensus heerst is er het probleem om de individuele bijdrage van de docent en de teambijdrage te scheiden. Sommige doelen kunnen alleen bereikt worden als docenten samenwerken, bijvoorbeeld het terugdringen van een spijbelprobleem op school. Een ander probleem betreft de losse relatie tussen inspanningen van docenten en leerlingprestaties. Er is nog geen blauwdruk voor goed onderwijzen; de benodigde technieken om de doelen te bereiken zijn niet duidelijk. Murmane en Cohen (1986) komen tot de conclusie dat het mislukken van de meeste plannen te maken heeft met de aard van het werk van de docent. Het onderwijzen voldoet volgens hen niet aan elementaire condities voor prestatiegerichte betaling.

In het geval de prestatiemeting op basis van een beoordeling door de schoolleider tot stand komt, treden vooral psychologische effecten op als docenten lager beoordeeld worden dan ze zelf hadden verwacht. Zo blijkt bijvoorbeeld dat docenten denken bij een latere beoordeling nooit lager te kunnen scoren dan bij een eerdere beoordeling. Bij onvrede over de resultaten kan gemakkelijk een vorm van onverschilligheid optreden met alle ongewenste gevolgen van dien. Een ander aspect is de veranderende rol van de directeur. Het lage moreel en de bestuurlijke problemen zijn dus vooral oorzaak van het mislukken van deze vorm van prestatiebeloning (Murmane & Cohen, 1986).

Firestone (1991) onderzocht door middel van case studies het verschil in effecten op de motivatie van docenten tussen een systeem van prestatiebeloning en een systeem van taakverbreding. Bij prestatiebeloning krijgen docenten meer geld omdat ze hetzelfde werk beter doen dan andere docenten. Bij taakverbreding worden sommige docenten extra betaald om meer en ander werk te doen. De conclusie is dat taakverbreding meer dan prestatiebeloning de motivatie van docenten vergroot. Bovendien verrijkt taakverbreding het dagelijks werk van de docent, terwijl prestatiebeloning dit juist standaardiseert. De motivatie van docenten wordt vooral vergroot door intrinsieke prikkels. Deze worden meer opgewekt door taakverbreding. Geld is dus niet direct van invloed op de verhoging van de motivatie van docenten, er kan echter wel intrinsieke motivatie mee worden 'gekocht'. Ook Firestone (1991) vond grote barrières voor invoering van individuele prestatiebeloning: docenten stellen vragen bij de eerlijkheid van de selectie, bij de nauwkeurigheid van de prestatievaststelling en bij de rechtvaardigheid van de verdeling. Tevens zou de collegialiteit een factor van betekenis vormen, met name omdat docenten zich geremd voelen om zich te onderscheiden van collega's.

Samenvattend kunnen we stellen dat systemen van prestatiebeloning van docenten sec in het onderwijs zelden succesvol zijn gebleken. Individuele beloningssystemen kunnen remmend werken op de teamarbeid die nodig is om een produktief schoolklimaat te ontwikkelen en te onderhouden (cf. Richards & Sheu, 1992).

1.2.2 Ervaringen met rendementsbekostiging

De Nederlandse overheid overweegt vooralsnog geen bonussen of kortingen op het niveau van de individuele docent, maar op het niveau van de school. Het Middelbaar Beroepsonderwijs wordt sinds 1991 geconfronteerd met lichte bekostigingsprikkels, doordat de scholen gekort worden voor studenten die de eindstreep niet of vertraagd halen. Over het effect van deze bekostigingsprikkel op het rendement zijn nog geen gegevens voorhanden. Wel is bekend dat binnen het Middelbaar Beroepsonderwijs zelf maatregelen worden getroffen die gericht zijn op rendementsverbetering, zoals intensivering van de leerlingbegeleiding, het opzetten van studievoortgangsregistratie en het vormen van docententeams voor de optimalisering van de leerlingbegeleiding, de -beoordeling en de -selectie (Brandsma & Thijsen, 1993).

Een goed voorbeeld van een outputgericht beloningssysteem in de Verenigde Staten is het 'South Carolina School Incentive Reward Program'. Het betreft hier een bekostigingssysteem, waarbij scholen met buitengewoon goede resultaten (im)materieel worden beloond. Men hoopt hiermee scholen te motiveren om te streven naar het behalen van bepaalde prestatiestandaarden. Deze standaarden zijn gedefinieerd op outputcriteria betreffende schoolprestaties (i.c. leerlingprestaties, gemeten als het verschil tussen de verwachte en de gerealiseerde prestatie en aanwezigheidspercentages van leerlingen en docenten). In het systeem wordt rekening gehouden met de leerlingpopulatie door scholen in te delen in groepen met een vergelijkbare populatie voor wat betreft achtergrondkenmerken die van belang worden geacht voor de te realiseren prestatie. Scholen worden ingedeeld op basis van het percentage leerlingen dat in aanmerking komt voor een gratis lunch (i.e. een indicator voor onderwijsachterstand), het percentage leerlingen dat in aanmerking komt voor een lunch tegen gereduceerd tarief, het gemiddeld aantal jaren onderwijservaring van de docenten en het percentage leerlingen dat op of boven de standaard voor een bepaalde leesvaardigheidstest scoort. Scholen 'wedijveren' dan ook binnen de eigen groep om de beschikbare bonussen. Scholen die wel vooruitgang boeken, maar nog niet in aanmerking komen voor een bonus komen in aanmerking voor een eervolle vermelding. In het betreffende bekostigingssysteem is ten slotte voorzien in sancties op frauderen met de leerlingprestaties en zijn voorschriften opgenomen met betrekking tot besteding van de bonus. Deze mag niet worden besteed aan beloning van docenten, maar moet ingezet worden voor verbetering van het onderwijsleerproces.

In de staat Georgia is een omvattend evaluatieprogramma ten behoeve van kwaliteitsbewaking van het onderwijs in gang gezet, waarbij de mogelijkheid van financiële prikkels wettelijk is verankerd. Het is een evaluatiesysteem met bekostigingprikkels, waarin ook scholen met dezelfde populaties met elkaar worden vergeleken (Shah & Rogers, 1991).

De ervaringen met het bekostigingssysteem in South Carolina zijn bemoedigend. Richards en Sheu (1992) concluderen na een analyse dat de leerlingprestaties in bescheiden mate zijn verbeterd, maar vonden geen toename in de presentie van docenten en leerlingen. Door de combinatie van de scores voor vooruitgang met de groepsindeling op achtergrondkenmerken van de leerlingpopulatie, blijkt de eigenstandige invloed van dergelijke achtergrondkenmerken op de verdeling van de financiële prikkels gereduceerd te worden. De ervaringen ten aanzien van de medewerking van de scholen zijn goed. Het systeem in Georgia bevindt zich nog in een ontwikkelingsfase.

1.3 Rendementsbekostiging

In Nederland is het veld van onderwijsfinanciering voortdurend in beweging. Het denken over rendementsbekostiging is daarbij een tamelijk nieuw fenomeen. De huidige bekostiging van het HAVO en het VWO geschiedt op normbasis voor de materiële uitgaven en op declaratiebasis voor de personele uitgaven middels het formatiebudgetsysteem. Het aantal ingeschreven leerlingen fungeert als maatstaf voor de vaststelling van de formatie, waarvan de kosten door de overheid worden vergoed. Het formatiebudgetsysteem beoogt scholen enige zelfstandigheid te geven om een beter personeelsbeleid te voeren en de onderwijskundige organisatie te verbeteren. De werkelijke beleidsruimte zal echter afhangen van het natuurlijke verloop onder het personeel en de ontwikkeling van het aantal leerlingen (Van der Bij et al., 1993).

In de huidige bekostiging zit volgens het

Ministerie van Onderwijs en Wetenschappen (1991) geen druk op scholen om leerlingen goed te selecteren, ze goed te begeleiden of om uitval tegen te gaan. HAVO- en VWO-scholen hebben er momenteel financieel gezien belang bij om zoveel mogelijk leerlingen binnen de school te houden. Zo is het voor HAVO-scholen bijvoorbeeld niet interessant om na het derde leerjaar leerlingen door te verwijzen naar het MBO.

Het denken over een vorm van outputfinanciering, waarbij scholen worden beloond voor kwaliteit, i.c. rendementsbekostiging, past goed in de nieuwe besturingsfilosofie van de overheid, waarin scholen meer autonomie krijgen en achteraf verantwoordelijk worden gesteld voor de geleverde kwaliteit. Door de overheid zijn in dit verband enkele alternatieve bekostigingsscenario's[2] ontwikkeld. Hierin wordt uitgegaan van een financiering die gedeeltelijk wordt gebaseerd op de aanwezigheid van de leerlingen en gedeeltelijk op het rendement van de bovenbouw van de school. Dit rendement van de bovenbouw wordt berekend als de ratio van het aantal geslaagden plus het aantal bevorderden op het aantal aanwezige leerlingen. Een bepaald percentage van de bekostiging wordt ingehouden en, met een bonus, uitgekeerd voor gediplomeerde leerlingen. In de scenario's kan de omvang van de korting of de bonus worden gevariëerd. In een eerste scenario is de korting relatief gering, namelijk 5%. Per schooljaar van de bovenbouw wordt voor iedere leerling 5% ingehouden. Deze korting kunnen scholen terugverdienen voor die leerlingen die in één keer het eindexamen halen en gediplomeerd de school verlaten. Bovenop dit terugverdienen van de korting, wordt voor elke geslaagde leerling een kleine bonus uitgekeerd. In een tweede en derde scenario zijn de kortingen en bonussen respectievelijk drie keer en tien keer zo groot als in het eerste scenario; de sterkte van de rendementsprikkel neemt dus sterk toe. In deze drie scenario's worden de CUMI-leerlingen nog volledig op basis van aanwezigheid gefinancierd. In een vierde scenario is een extra hoge bonus voorzien voor CUMI-leerlingen die in één keer en zonder vertraging het diploma halen.

Er wordt in deze bekostigingsscenario's een kwantitatieve invulling aan het begrip rendement gegeven. Het gaat om het aantal leerlingen dat het diploma haalt en om de efficiëntie van de loopbanen door de bovenbouw, waarbij efficiënt duidt op het behalen van het diploma binnen de nominale studieduur. Dit impliceert dat uitgegaan wordt van een gegeven kwaliteit van de output van het onderwijsproces, weerspiegeld in de exameneisen. In de scenario's wordt op drie manieren rekening gehouden met de samenstelling van de leerlingpopulatie, zij het in beperkte mate. Op de eerste plaats worden in de eerste drie scenario's de CUMI-leerlingen buiten beschouwing gelaten in de rendementsberekening. Het korting- en bonussysteem is op hen niet van toepassing. Voor het VWO wordt tevens rekening gehouden met de samenstelling van de leerlingpopulatie door in de scenario's onderscheid te maken tussen aanwezige leerlingen met en zonder HAVO-diploma, waarbij de laatsten, gezien de door hen gemaakte omweg, in mindere mate bekostigd worden. Een andere poging om de samenstelling van de leerlingpopulatie te verdisconteren in de bekostigingsberekening komt tot uitdrukking in het vierde alternatieve scenario, waarin een extra bonus wordt uitgekeerd voor geslaagde leerlingen uit de groep culturele minderheden.

Toepassing van deze vorm van rendementsbekostiging leidt tot budgetkortingen voor scholen die beneden het landelijk gemiddelde rendement scoren en tot bonussen voor scholen die boven dit gemiddelde scoren. Hoewel het idee is dat rendementsbekostiging uiteindelijk tot besparingen op macroniveau leidt, kunnen de budgettaire consequenties thans nog niet worden vastgesteld.

1.4 Strategisch beleid en strategisch gedrag

Vanuit haar zorg voor de kwaliteit van het onderwijs en de beheersing van de overheidsuitgaven overweegt de overheid aldus de bekostiging van HAVO- en VWO-scholen gedeeltelijk afhankelijk te stellen van het gerealiseerde rendement (Ministerie van Onderwijs en Wetenschappen, 1991). De verwachting is daarbij dat de bekostigingsprikkel scholen zal bewegen tot het voeren van een beleid gericht op rendementsverbetering.

Scholen hebben hierbij te maken met drie beperkende factoren. Allereerst mag, zoals gezegd, de mate waarin scholen via beleidsmaatregelen het rendement kunnen verbeteren niet

overschat worden. Het rendement wordt grotendeels bepaald door de samenstelling van de leerlingpopulatie, die door scholen maar in zeer beperkte mate te beïnvloeden is. Bovendien is voor het treffen van onderwijsinhoudelijke en -organisatorische maatregelen ter verbetering van het rendement nog geen draaiboek voorhanden. Op de tweede plaats hebben scholen te maken met beperkingen in het kader van hun beleidsvoering die voortvloeien uit de wettelijke voorschriften ten aanzien van de inrichting van het onderwijs, zoals de bekostigingsvoorwaarden en de deugdelijkheidseisen. De beleidsruimte van scholen in het algemeen is enigszins toegenomen ten gevolge van de deregulering, de schaalvergroting en de invoering van het formatie-budgetsysteem. De (resterende) vrijheid van inrichting die scholen wettelijk gezien hebben, behelst zowel de onderwijsinhoudelijke vormgeving als de beheersmatige vormgeving. Met name op het gebied van de toelating van leerlingen, de keuze van de leermiddelen en de aanstelling van onderwijsgevenden heeft het bijzonder onderwijs hierin iets meer beleidsruimte dan het openbaar onderwijs. Ten slotte kan men denken aan de mate waarin de organisatorische inrichting van de school haar in staat stelt beleid te voeren. We spreken dan over de relatie tussen het beleidsvoerend vermogen van de school en de organisatiestructuur (cf. Sleegers, 1991).

Ter onderscheiding van mogelijk beleid dat scholen kunnen gaan voeren als reactie op de rendementsbekostiging waarmee ze worden geconfronteerd worden in het hierna beschreven onderzoek de termen strategisch beleid en strategisch gedrag, of ontwijkingsgedrag gehanteerd. Bij strategisch **beleid** hebben we dan de beleidsvoering voor ogen die enerzijds gericht is op de verbetering van het rendement van de eigen school en anderzijds bijdraagt aan de realisatie van de overheidsdoelen, namelijk het in samenhang vergroten van het kwalitatief en het kwantitatief rendement van de bovenbouw van het voortgezet onderwijs. Een belangrijk criterium is dat deze maatregelen zodanig ten goede komen aan de leerlingen dat uiteindelijk het niveau van de leerprestaties wordt verhoogd. Strategisch beleid verwijst aldus naar maatregelen die leiden tot de gewenste effecten voor het voortgezet onderwijs. De veronderstelling is dat dit beleid fundamentele beslissingen omvat en een lange-termijn karakter draagt. De resultaten kunnen pas na enige tijd verwacht worden. Beleid onder de noemer strategisch beleid kan zowel betrekking hebben op de beheersmatige als op de onderwijsinhoudelijke aspecten van de school. Te denken valt hierbij aan de begeleiding van leerlingen, de beoordeling en de selectie van leerlingen, de coördinatie tussen docententeams met het oog op optimale afstemming en veranderingen in het schoolwerkplan.

Bij strategisch **gedrag** hebben we de beleidsvoering voor ogen die weliswaar gericht is op rendementsverbetering van de eigen school, doch niet bijdraagt aan de realisatie van de overheidsdoelen ten aanzien van de verbetering van zowel het kwantitatieve als het kwalitatieve rendement van de bovenbouw van het voortgezet onderwijs. Het gaat daarbij om beleidsmaatregelen die primair gericht zijn op handhaving, dan wel uitbreiding van de bekostiging voor de eigen school, c.q. budgetmaximalisatie. Het belang van de leerling staat daarbij niet voorop. Strategisch gedrag verwijst aldus naar beleidsmaatregelen die leiden tot ongewenste neveneffecten. We kunnen hier te maken hebben met ad hoc beslissingen als reactie op een wellicht tegenvallende bekostiging. Voorbeelden hiervan zijn ingrepen in het schoolonderzoek (verlagen van niveau door vereenvoudiging van opgaven, het laten leren voor de toets of het achteraf aanpassen van de norm) of een tijdelijke verscherping van de doorstroming van de onderbouw naar de bovenbouw (bijstellen overgangsnormen, 'bespreekzone' voor twijfelgevallen versmallen of verzwaren proefwerken). Illustratief in dit verband is de geconstateerde stijging van het aantal zittenblijvers in HAVO-4 en VWO-5 sinds de invoering van de 10%-gezaktenmaatregel (Van der Bij et al., 1993). Er kan echter ook sprake zijn van een weloverwogen lange-termijnbeleid, waarbij men tracht via een aannamebeleid zodanige invloed op de kenmerken van de leerlingpopulatie uit te oefenen dat een rendementsverbetering op termijn in het verschiet ligt.

De verwachting is nu dat scholen eerder geneigd zijn strategisch gedrag te vertonen als ze worden geconfronteerd met een korting op basis van een tegenvallend rendement. In zo'n

situatie zijn scholen gedwongen snel (en met minder faciliteiten) hun rendement te verhogen om hun bekostiging weer op peil te krijgen. Van strategische beleidsmaatregelen is over het algemeen pas op termijn resultaat te verwachten, terwijl deze beleidsvoering daarnaast in veel gevallen extra investeringen vraagt. Men zou kunnen zeggen dat scholen die een bonus ontvangen het zich kunnen permitteren om strategische beleidsmaatregelen te treffen.

2 Methode

2.1 Steekproef

De kernvraag van het onderzoek betreft het oordeel van schoolleiders over rendementsbekostiging en de mogelijke beleidsreacties binnen de scholen.

Voor de populatie HAVO- en VWO-scholen zijn met behulp van de integrale leerlingtellingen het rendement en de financiële consequenties van de vier bekostigingsscenario's berekend. Vervolgens is een gestratificeerde steekproef van 56 scholen getrokken met als stratificatiecriteria denominatie, schooltype, urbanisatiegraad, samenstelling scholen (gemeenschap), rendement en percentage CUMI-leerlingen. Er moesten 94 scholen worden benaderd om zo dicht mogelijk bij het geplande aantal van 56 scholen te komen. Uiteindelijk hebben 53 scholen hun medewerking aan het onderzoek verleend (een responspercentage van 56%). Het ging daarbij om 25 HAVO-scholen (respons: 50%), 21 VWO-scholen (respons: 65.6%) en 7 scholen met een hoog percentage CUMI-leerlingen (respons: 48.3%). De kenmerken van de steekproef in vergelijking met de populatie worden in Tabel 1 weergegeven.

Er is in de tabel te zien dat er sprake is van onder- en oververtegenwoordiging in de steekproef ten opzichte van de populatie. Deze verstoringen komen voort uit de gekozen werkwijze inzake selectie van de scholen, waarbij gestreefd is naar een optimale stratificatie, dat wil zeggen één school per stratificatiecel, zodat zinvolle vergelijkingen tussen diverse groepen scholen kunnen worden gemaakt. Deze spreiding over de cellen was niet altijd mogelijk; de scholen met veel CUMI-leerlingen zijn meestal openbaar en gehuisvest in de grote steden en bijzondere scholen met een zeer laag rendement komen nauwelijks voor.

2.2 Dataverzameling en instrumenten

De data zijn verzameld met behulp van twee instrumenten. In een schriftelijke vragenlijst zijn voornamelijk (administratieve) gegevens verzameld die noodzakelijk zijn om het feitelijk rendement van de responderende scholen te berekenen en de situatie rond het onderwijskundige beleid in kaart te brengen. De schriftelijke vragenronde werd gevolgd door een gestructureerd interview met conrectoren van respectievelijk de HAVO- en VWO-afdeling, waarin het oordeel over rendementsbekostiging, de beleidsmaatregelen die scholen zouden treffen

Tabel 1
Vergelijking kenmerken steekproef met kenmerken populatie

	HAVO (*n*=25)		VWO (*n*=21)		CUMI (*n*=7)	
	steekpr.	pop.	steekpr.	pop.	steekpr.	pop.
	%	%	%	%	%	%
Denominatie						
bijzonder	52.0	76.0	42.9	71.4	14.3	33.3
openbaar	48.0	24.0	57.1	28.6	85.7	66.7
Schooltype						
LBO-MAVO-HAVO-VWO	16.0	6.5	0	4.9	0	17.2
MAVO-HAVO-VWO	44.0	55.8	42.9	49.4	85.7	70.0
HAVO-VWO	40.0	37.7	47.6	35.1	14.3	12.5
VWO			9.5	10.2	0	0
Urbanisatiegraad						
< 100.000 inw.	48.0	69.9	38.1	67.9	0	29.7
≥ 100.000 inw.	52.0	30.1	61.9	32.1	100	70.3

bij invoering van rendementsbekostiging en de beleidsvorming binnen de school aan de orde kwamen. Om de respondent in de gelegenheid te stellen intern overleg te voeren over de mogelijke reacties van de school op invoering van een vorm van rendementsbekostiging zijn de onderwerpen die in het interview aan de orde komen vooraf aan de respondent meegedeeld.

De respondenten is een elftal beleidsterreinen voorgelegd, met de vraag aan te geven op welke van deze terreinen maatregelen prioriteit zouden krijgen indien een vorm van rendementsbekostiging zou worden ingevoerd. Daarnaast is de respondenten een groot aantal items voorgelegd, welke concrete maatregelen representeerden, met de vraag aan te geven hoe groot de kans is dat deze maatregel zou worden getroffen indien een vorm van rendementsbekostiging wordt ingevoerd.

Om de mogelijke beleidsgevolgen van invoering van de vier bekostigingsmodellen te kunnen weergeven in termen van strategisch beleid en strategisch gedrag zijn twee schalen geconstrueerd. De constructie van de eerste schaal verliep als volgt. Eerst zijn de beleidsprioriteiten van scholen bij invoering van rendementsbekostiging via het interview geïnventariseerd. Een a-priori ordening van beleidsprioriteiten leidt tot het continuum:

strategisch beleid ▲
- a - optimaliseren leerlingbegeleiding
- b - optimaliseren leerlingbeoordeling
- c - optimaliseren selectie- en determinatie
- d - optimaliseren functioneren organisatie
- e - stimuleren (vak)inhoudelijke verbeteringen
- f - verscherpen personeelsbeleid
- g - veranderen schoolwerkplan
- h - verscherpen doorstroombeleid
- i - verscherpen doublurebeleid
- j - verscherpen aannamebeleid
- k - versoepelen bewaking kwalitatief niveau

strategisch gedrag ▼

Om te kijken of deze ordening hout snijdt en om vast te stellen waar scholen zich op dit continuum bevinden is stochastische ontvouwing als analyse-techniek toegepast (Van Schuur, 1984). Na analyse bleek dat drie beleidsterreinen niet op het te ontvouwen continuum geplaatst konden worden. De uiteindelijke schaal en de statistische gegevens staan weergegeven in Tabel 2.

De *H(I)*-coëfficiënt geeft aan in hoeverre een maatregel schaalbaar is. De coëfficiënt loopt van 0 (niet schaalbaar) tot 1 (perfect schaalbaar). In het algemeen duidt een waarde van .30 op voldoende en .40 op redelijke schaalbaarheid. Daarnaast is de *T*-waarde vermeld, die wanneer groter dan 1.96, aangeeft dat de geschatte *H(I)*-waarde statistisch van 0 (niet schaalbaar) afwijkt met 5% kans op een verkeerde beslissing (in werkelijkheid was de maatregel niet schaalbaar). De totale schaal is voldoende tot redelijk homogeen; het verwijderen van item h (verscherpen doorstroombe-

Tabel 2
Ontvouwingsgegevens

				H(I)	*T(H)*
strategisch beleid	▲				
		d	- optimaliseren functioneren organisatie	.62	3.04
		a	- optimaliseren leerlingbegeleiding	.41	4.21
		e	- stimuleren (vak)inhoud. verbeteringen	.42	5.34
		g	- veranderen schoolwerkplan	.40	5.51
		c	- optimal. selectie- en determinatiebeleid	.31	4.29
		j	- verscherpen aannamebeleid	.34	3.73
		i	- verscherpen doublurebeleid	.39	3.20
strategisch gedrag		h	- verscherpen doorstroombeleid	.22	1.39
	▼				
Totale schaal				.38	4.79

leid) leidt niet tot een verbetering, zodat dit item ondanks de matige itemkarakteristiek toch gehandhaafd blijft.

Naast het aangeven van beleidsprioriteiten is scholen gevraagd naar concrete maatregelen die genomen zouden worden bij invoering van rendementsbekostiging. Hiervoor is een Likertschaal ontwikkeld voor de dimensie van strategisch beleid tot strategisch gedrag. De betrouwbaarheid van deze schaal is goed (Cronbach's $\alpha = 0.78$).

De verzamelde data voor het totale onderzoek zijn geanalyseerd met behulp van rechte tellingen, variantie-, correlationele analyses en schaaltechnieken.

3 Resultaten

Als resultaat van de rendementsanalyses komt naar voren dat rendementsbekostiging tot een hogere financiering van bepaalde groepen scholen leidt. Zo kunnen bijzondere scholen, grote scholen en scholen in kleine steden gemiddeld een bonus tegemoet zien. Daarentegen krijgen openbare scholen, kleine scholen en scholen in de grote steden met een korting te maken. De antwoorden op de onderzoeksvragen naar het oordeel over rendementsbekostiging, de beleidsreacties en de veranderingen in beleid komen hierna aan de orde.

3.1 Oordeel over rendementsbekostiging

Ruim driekwart van de scholen blijkt op de hoogte te zijn van de voornemens van het Ministerie om het terrein van rendementsbekostiging te gaan verkennen. Over het algemeen oordelen de scholen uitgesproken negatief over rendementsbekostiging, zoals te zien is in Tabel 3.

Respectievelijk 64.2% en 60.4% van de scholen is tegen rendementsbekostiging in het algemeen dan wel tegen rendementsbekostiging voor alleen de bovenbouw. Er worden met name negatieve gevolgen voor de leerlingen gevreesd. Zwakke leerlingen en laatbloeiers zullen minder kansen krijgen een diploma te halen, zo oordelen de schoolleiders. Het minst negatief wordt geoordeeld over een vorm van rendementsbekostiging waarbij een bonus voor geslaagden uit achterstandsgroepen, zoals culturele minderheden, wordt uitgekeerd (ruim 27% oordeelt positief en circa 26% heeft een neutraal oordeel). Het begeleiden van leerlingen in achterstandssituaties vraagt een extra prestatie van scholen en deze mag beloond worden. Het oordeel over rendementsbekostiging hangt niet samen met het huidige rendement van de school, de denominatie, de urbanisatiegraad of het schooltype. Openbare scholen, scholen in grote steden, HAVO-scholen of scholen met een hoog rendement verschillen dus niet in hun oordeel over rendementsbekostiging in vergelijking met bijzondere scholen, scholen op het platteland, VWO-scholen of scholen met een laag rendement.

3.2 Beleidsreacties bij rendementsbekostiging

Zowel de algemene beleidsterreinen als de concrete maatregelen kunnen worden afgebeeld op de dimensie van strategisch gedrag naar strategisch beleid, respectievelijk met behulp van de ontvouwingschaal en de Likertschaal. Er is vervolgens een variantie-analyse uitgevoerd van de scores op de ontvouwingsschaal en van de scores op de Likertschaal met de variabelen denominatie, urbanisatiegraad, rendement, schooltype, onderwijssoort en oordeel over rendementsbekostiging. Alleen de gevonden significante verbanden zijn weergegeven in

Tabel 3
Oordeel over rendementsbekostiging

	negatief %	neutraal %	positief %	n
oordeel over rendementsbekostiging in het algemeen	64.2	20.8	15.0	53
oordeel over rendements bekostiging voor de bovenbouw	60.4	24.5	15.1	53
oordeel over rendements bekostiging met bonus voor achterstandsgroepen	46.5	25.8	27.6	48

Tabel 4
Strategisch beleid, strategisch gedrag naar denominatie, rendement en onderwijssoort

	Algemeen beleid		Concrete maatregelen		
	Gem.	*SD*	*Gem.*	*SD*	*n*
Denominatie[1]					
Bijzonder			3.17	0.47	23
Openbaar			2.97	0.41	30
Rendement[2]					
laag	3.82	1.33	2.90	0.43	17
midden	4.23	1.60	2.98	0.42	17
hoog	5.18	1.69	3.26	0.41	19
Onderwijssoort[3]					
HAVO			3.17	0.36	28
VWO			2.93	0.50	25

1 p=0.10; Eta^2= 0.05
2 p=0.03; Eta^2= 0.13
3 p=0.04; Eta^2= 0.08

Tabel 4. De schaalscores voor algemeen beleid lopen daarbij van 0 (strategisch gedrag) tot 7 (strategisch beleid). De schaalscores voor de concrete maatregelen lopen van 1 tot 5, waarbij 1 duidt op strategisch gedrag en 5 duidt op strategisch beleid.

Ten aanzien van de prioritering van de algemene beleidsterreinen blijkt dat scholen met een hoog rendement relatief meer geneigd zijn te kiezen voor terreinen die tot het strategisch beleid kunnen worden gerekend, dan scholen met een laag rendement. Deze uitkomsten worden bevestigd door de positie van de scholen op de Likertschaal, die is geconstrueerd op basis van hun antwoorden ten aanzien van de concrete maatregelen.

Het blijkt dat de bijzondere scholen, de scholen met een hoog rendement en de HAVO-scholen iets meer geneigd zijn tot strategisch beleid dan de openbare scholen, de scholen met een laag rendement en de VWO-scholen. Hierbij moet echter wel worden aangetekend dat scholen over het algemeen meer neigen naar strategisch beleid dan naar strategisch gedrag. Extreme vormen van strategisch gedrag, zoals versoepeling van de bewaking van het kwalitatief niveau (bijvoorbeeld verlagen van de moeilijkheidsgraad van de schoolonderzoeken), worden amper of niet als beleidsoptie genoemd.

Wat betreft de concrete maatregelen, blijkt dat meer dan 50% van de scholen het zeer waarschijnlijk acht dat bij invoering van een vorm van rendementsbekostiging de volgende maatregelen zullen worden getroffen:

1 creëren van meer bijscholingsmogelijkheden voor docenten;
2 invoeren dan wel automatiseren van een leerlingvolgsysteem;
3 stoppen van meer tijd per leerling in studiebegeleiding;
4 voorzichtiger adviseren bij twijfel over schooltypekeuze;
5 vaker doorstroming van HAVO-3 naar het MBO adviseren;
6 sneller de afstroom naar een lager schooltype adviseren.

3.3 Veranderingen in beleid

De maatregelen die de scholen zeggen te gaan nemen zijn, teneinde inzicht te krijgen in de veranderingen in beleid, vergeleken met het huidige beleid dat de school voert. Tabel 5 toont dat scholen meer geneigd zijn strategische beleidsmaatregelen ter verbetering van hun rendement te treffen naarmate het organisatie- en onderwijskundige beleid op de corresponderende terreinen in mindere mate ontwikkeld is.

Dit komt meer specifiek tot uitdrukking bij aspecten van de leerlingbegeleiding, de leerlingbeoordeling en de schoolorganisatie. Hoe minder faciliteiten men in de onderbouw voor de leerlingbegeleiding kent, des te eerder kiest men voor beleidsterreinen en concrete maatregelen die kunnen worden gerekend tot strategisch beleid. Verder geldt dat scholen verhoudingsgewijs meer kiezen voor de strategische

Tabel 5
Samenhang tussen aspecten van huidige en toekomstige schoolsituatie

	KEUZE BELEIDSTERREINEN	CONCRETE MAATREGELEN
AANNAME		
toelatingscriteria onderbouw	−.10	*.16*
toelatingscriteria bovenbouw HAVO	*−.18*	−.01
toelatingscriteria bovenbouw VWO	−.17	*−.27*
LEERLINGBEGELEIDING		
voorzieningen onderbouw	**−.27**	**−.23**
voorzieningen bovenbouw	−.09	−.01
aantal uren onderbouw	.02	*.16*
aantal uren bovenbouw	−.02	.12
LEERLINGBEOORDELING		
beoordelingssysteem	**−.25**	−.08
DOORSTROMING		
criteria advies na brugklas	*−.13*	−.15
SCHOOLORGANISATIE		
aantal functioneringsgesprekken	−.03	−.06
verzuimmaatregelen	**−.33**	*−.15*
SELECTIE/DETERMINATIE		
advisering vakkenpakketkeuze	.01	.13
BESTUURLIJKE BETROKKENHEID		
organisatiebeleid	.01	.11
onderwijskundig beleid	*−.13*	−.13
PRESTATIEGERICHTHEID	−.10	.02

vet gedrukte correlaties zijn significant bij $\alpha < .10$
cursief gedrukte correlaties zijn significant bij $\alpha < .20$

beleidsterreinen, naarmate ze in de huidige situatie minder maatregelen kennen om ongeoorloofd verzuim van leerlingen tegen te gaan en naarmate hun beoordelingssysteem minder systematisch is.

4 Discussie

In het voorgaande is ingegaan op het onderwijseconomische input-outputonderzoek, het schooleffectiviteitsonderzoek en zijn ervaringen met bekostigingsprikkels alsmede enkele resultaten van een onderzoek naar oordelen over rendementsbekostiging in de bovenbouw van het HAVO/VWO beschreven. Aan het voornemen om rendementsbekostiging in te voeren ligt de assumptie ten grondslag dat scholen in staat moeten worden geacht via beleidsmaatregelen hun rendement te verbeteren. Uit reeds eerder uitgevoerd onderzoek blijkt dat deze aanname allerminst sterk is. Uit de onderzoeksliteratuur weten we immers dat het rendement van scholen vooral een resultante van de instroom is. Scholen kunnen hun rendement via organisatorische- en onderwijsinhoudelijke beleidsingrepen maar ten dele manipuleren. Daar komt bij dat het allerminst duidelijk is welke maatregelen nu precies tot rendementsverbetering kunnen leiden. Desalniettemin hebben scholen kennelijk wel vertrouwen in de mogelijkheden om hun rendement te manipuleren; zij reageren, met andere woorden, alsof er een produktiefunctie bestaat. Het lijkt er op dat bekostigingsprikkels op zich een effectief middel zouden kunnen zijn om scholen te stimuleren beleidsaanpassingen te doen. Zowel ervaringen met prestatiebeloning van docenten als met rendementsbekostiging in het buitenland, alsmede de te verwachten beleidsreacties van HAVO- en VWO-scholen tonen aan dat docenten en scholen gevoelig zijn voor deze prikkel en hun gedrag aanpassen. Het verbinden van financiële consequenties aan een

systeem van outputcontrole is overigens een tamelijk drastische maatregel. Gelet op de noodzaak om met uitermate verfijnde scenario's te werken en gelet op het bestaan van vele valkuilen die bij een gebrekkige implementatie kunnen optreden, zou het aanbeveling verdienen om eerst wat terughoudender middelen in te zetten, zoals het publiceren van rendementsgegevens van scholen. Wellicht kan van zo'n middel ook de gewenste stimulans uitgaan. Zolang er echter een zeer geringe empirische bewijslast bestaat voor de gedachte dat het rendement van scholen een functie is van een bepaalde organisatorische en onderwijsinhoudelijke setting, kan de overheid ons inziens eigenlijk niet volstaan met outputcontrole. Aandacht van de overheid voor de mogelijk relevante proceskenmerken kan bijdragen aan het ontstaan en verspreiden van meer inzicht in de produktiefunctie van het onderwijs.

Een andere belangrijke conclusie betreft de reeds genoemde invloed van de kenmerken van de leerlingpopulatie op het rendement. Uit de rendementsanalyses komt naar voren dat rendementsbekostiging tot kortingen leidt voor openbare scholen, kleine scholen en scholen in grotere steden. Zoals de scenario's er nu uitzien, zullen ze een onrechtvaardige uitwerking hebben. Men dient slechts te bedenken dat grote openbare scholen in grote steden meer dan andere scholen risicoleerlingen huisvesten. Het gaat hierbij niet alleen om allochtone leerlingen, maar ook om leerlingen uit de lagere sociaal-economische milieus. Ook de rode draad door de antwoorden van de respondenten, dat betalen op basis van output, betalen op basis van input betekent, kan in dit verband verontrustend worden genoemd. De mogelijkheden van scholen om hun instroom te reguleren zijn in het algemeen gering, maar voor bijzondere scholen en scholen buiten de grote steden relatief groter. Het feit dat in de thans ontwikkelde scenario's te weinig rekening gehouden wordt met verschillen in leerlingpopulaties kan als de belangrijkste implementatiedrempel worden opgevat. Het laten wedijveren van scholen in groepen van scholen met een vergelijkbare leerlingpopulatie is daartoe elders een werkbare oplossing gebleken. Ook uit andere ervaringen met outputgerichte evaluatie- en bekostingssystemen (cf. May, 1990; Shah & Rogers, 1991; Tesh, 1991; Richards & Sheu, 1992) weten we dat de medewerking van scholen en docenten slechts gewaarborgd is, als het systeem als solide, fair en duidelijk wordt ervaren. Dat betekent in ieder geval dat scholen slechts verantwoordelijk mogen worden gesteld voor factoren die zij in enige mate kunnen beheersen. Voor een succesvolle implementatie is het verder belangrijk dat de betrokkenen de bonus de moeite waard vinden om zich er extra voor in te spannen, dat de criteria communicatief zijn en dat het systeem niet zodanig uitwerkt dat steeds dezelfde scholen winnen of verliezen; dit is frustrerend voor de motivatie.

Hoewel de vrees van de overheid voor strategisch, ofwel ontwijkingsgedrag op basis van de onderzoeksresultaten niet gegrond is, lijkt serieuze aandacht voor de randvoorwaarden van implementatie onvermijdelijk. Psychologische effecten en strategisch gedrag die tot tegengestelde resultaten kunnen leiden liggen wel degelijk op de loer bij de implementatie van een als onzorgvuldig, onduidelijk, of onrechtvaardig ervaren systeem van rendementsbekostiging.

Noten

1 Het onderzoek is uitgevoerd in opdracht van het Ministerie van Onderwijs en Wetenschappen bij het Onderzoekcentrum Toegepaste Onderwijskunde aan de Universiteit Twente; SVO-project 0524.

2 De hier beschreven mogelijke scenario's zijn ontwikkeld door het Ministerie van Onderwijs en Wetenschappen.

Literatuur

Amelsvoort, H.W.C.H. van, Bosker, R.J., & Brandsma, T.F. (1992). *Outputfinanciering: pro's en contra's.* Enschede: Universiteit Twente, OCTO.

Bij, J. van der, Noort, E.A. van, & Vrancken, P.H.J. (1993). *Kosten en baten van rendementsbekostiging in de bovenbouw van het AVO.* Den Haag: IOO.

Bosker, R.J., Guldemond, H., Hofman, R.H., & Hofman, W.H.A. (1989). *Kwaliteit in het voortgezet onderwijs.* Groningen: RION.

Brandsma, T. F., & Thijsen, J. A. (1993). *Flexibiliteit in de kwalificatievoorziening.* Enschede: Onderzoek Centrum Toegepaste Onderwijskunde (in voorbereiding).

Brempong, K. G., & Gyapong, A. O. (1991). Characteristics of education production functions: An application of Canonical Regression Analysis. *Economics of Education Review, 10*(1), 7-17.

Firestone, W. A. (1991). Merit pay and job enlargement as reforms: Incentives, implementation, and teacher response. *Educational Evaluation and Policy Analysis, 13,* 269-288.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24,* 1141-1177.

May, J. (1990). *Real world considerations in the development of an effective school incentive program.* (AERA-paper). Boston, SC: South Carolina Department of Education.

McCarty, T. A., & Brazer, H. E. (1990). On equalizing school expenditures. *Economics of Education Review, 9,* 251-264.

Ministerie van Financien (1991). *Rijksbegroting 1992.* Hoofdstuk VIII. Den Haag: SDU.

Ministerie van Onderwijs en Wetenschappen (1991). *Profiel tweede fase voortgezet onderwijs.* Den Haag: SDU.

Monk, D. H. (1991). *Education productivity research: An update and assessment of its role in education finance reform.* (AERA-paper). San Francisco: Maryland Bethesda.

Murmane, R. J., & Cohen, D. K. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. *Harvard Educational Review,* 56(1), 1-17.

Richards, C. E., & Sheu, T. M. (1992). The South Carolina school incentive reward program: A policy analysis. *Economics of Education Review, 11,* 71-86.

Scheerens, J., & Creemers, B. P. M. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research, 13,* 691-705.

Schuur, H. van (1984). *Structure in political beliefs, a new unfolding model with application to European party activists.* Amsterdam: CT Press.

Shah, J. G., & Rogers, W. (1991). *A system for statewide school accountability.* (AERA-paper). Georgia: Department of Education.

Sleegers, P. J. C. (1991). *School en beleidsvoering.* Academisch proefschrift, Katholieke Universiteit, Nijmegen.

Tesh, A. S. (1991). *A research-based attribute for school accountability.* (AERA-paper). Greensboro: University of North Carolina.

Walberg, H. J. (1992). Educational indicators for educational progress. In B. P. M. Creemers & W. Th. J. G. Hoeben (Eds.). *Indicatoren van onderwijseffectiviteit.* Groningen: ICO.

Webster, W. J. Mendro, R. L., & Almaquer, T. O. (1992). *Measuring the effects of schooling: expanded school effectiveness indices.* (AERA-paper). Dallas: Independent School District.

Manuscript aanvaard 19-4-1993

Auteurs

H. W. C. H. van Amelsvoort, R. J. Bosker en **T. F. Brandsma** zijn respectievelijk als universitair docent/medewerker onderzoek, universitair hoofddocent, en medewerker onderzoek/coördinerend projectleider verbonden aan de Vakgroep Onderwijsorganisatie- en management en het OCTO van de Universiteit Twente.

Adres: Universiteit Twente, OCTO/Vakgroep Onderwijsorganisatie en -management, Postbus 217, 7500 AE Enschede.

Abstract

Output financing in Dutch General Secondary Education

H. W. C. H. van Amelsvoort, R. J. Bosker & T. F. Brandsma. Pedagogische Studiën, 1993, 70, 283-297.

The Dutch Ministry of Education considers the introduction of a system of output financing for the superstructure of general secondary education (HAVO/VWO). Outputfinancing is perceived as an incentive for schools to improve both their quantitative and their qualitative rate of succes. However, the extent to which schools can manipulate their performances should not be overestimated. It is known that the output of schools is mainly determinated by the entrance characteristics of their pupils. Moreover, it is not clear which measures schools should take in order to improve their performances. Yet, the results from several studies in the USA as wells as a survey at 53 Dutch general secondary schools show that

schools do tend to take actions aimed at improving their output, once they are confronted with financial incentives. This survey did not give any evidence that schools would display evasive behavior, characterized by improving their rate of succes in favor of maximalisation their budgets and at the costs of their pupils' benefits.

Instructie in strategieën voor begrijpend lezen aan zwakke lezers

A. M. A. Walraven, P. Reitsma en E. J. Kappers

Samenvatting

In een interventie-onderzoek (pretest-posttest control group design) kregen zwakke lezers instructie in het gebruik van strategieën voor begrijpend lezen. Het doel was dat ze hun leesgedrag beter gingen plannen, bewaken en evalueren. De vraag was of deze instructie een positief effect zou hebben op prestaties in begrijpend lezen. Tevens werd nagegaan in hoeverre het aanleren van een strategie betreffende het activeren van voorkennis, effectief was. Aan het onderzoek namen 24 leerlingen uit het speciaal onderwijs deel (leeftijd 10-12 jaar). Twaalf leerlingen werd, in groepjes van drie, een aantal strategieën aangeleerd via directe instructie en 'reciprocal teaching'. Van deze groep kregen zes leerlingen een extra strategie aangeleerd: het activeren van voorkennis. Uit de resultaten blijkt dat de leerlingen uit de experimentele groep, in vergelijking met een controlegroep, meer kennis van strategieën krijgen en beter gaan presteren op toetsen voor begrijpend lezen: de Invultest en de Hoofdgedachtetest. Zowel instructie zonder als instructie inclusief activatie van voorkennis blijkt effectief.

Inleiding

Wanneer is een lezer een goede begrijpende lezer? In recente opvattingen over begrijpend lezen wordt de actieve, controlerende en regulerende rol van een goede lezer sterk benadrukt. Lezen wordt opgevat als een (inter)actief proces waarbij de lezer flexibel verschillende strategieën hanteert om een tekst goed te begrijpen. Zwakke lezers kunnen in vergelijking met goede lezers vooral als passief worden getypeerd: het zijn lezers die hun leesgedrag weinig of niet controleren en bijsturen (Bristow, 1985).

Binnen de benadering van lezen als interactief proces staat het begrip metacognitie centraal. Metacognitie kan worden opgevat als de kennis en controle die leerlingen hebben over hun eigen denk- en leeractiviteiten (Brown, 1980; Baker & Brown, 1984). Toegespitst op het proces van begrijpend lezen gaat het vooral om het controleren van het leesbegrip en het flexibel en adequaat plannen, sturen en evalueren van het leesproces (Baker & Brown, 1984; Forrest-Pressley & Gillies, 1983). Voor dit sturen en controleren van het leesgedrag is het van belang dat de lezer inhoudelijke kennis heeft van verschillende strategieën: de verklarende ('declarative') kennis. Vervolgens is het nodig om te weten hoe een bepaalde strategie toegepast moet worden: de procedurele ('procedural') kennis. Tenslotte is het voor een flexibel gebruik van strategieën van belang te weten wanneer een strategie wel en wanneer niet toegepast kan worden: de conditionele ('conditional') kennis (Paris, Cross & Lipson, 1984).

Instructie in begrijpend lezen richt zich binnen een dergelijke (meta-)cognitieve benadering vanzelfsprekend op het aanleren van strategieën die het leesproces ondersteunen en vergemakkelijken. Enkele interventiestudies hebben reeds aangetoond dat instructie in strategieën een positief effect kan hebben op kennis en gebruik van strategieën.

Een eerste voorbeeld daarvan is het werk van Duffy en collega's (Duffy, Roehler, Sivan, Rackliffe, Book, Meloth, Vavrus, Wesselman, Putnam & Bassari, 1987). In hun aanpak wordt getracht deelvaardigheden van begrijpend lezen, die veelal statisch zijn, te onderwijzen als strategieën. In vier fasen worden de vaardigheden als flexibele, aanpasbare en situatie afhankelijke activiteiten aangeboden, waarbij voordoen en hardop denken belangrijk zijn. Uit onderzoek bleek dat deze benadering succesvol was. Zwakke lezers uit de experimentele groep hadden na afloop meer kennis van strategieën en presteerden beter op een formatieve toets, vergeleken met leerlingen uit een controlegroep. Transfer naar een traditionele toets

voor begrijpend lezen bleek echter niet op te treden.

Een aanpak waarin metacognitieve strategieën een centrale plaats innemen, is het uitgebreide curriculum 'Informed Strategies for Learning' van Paris en collega's (Paris et. al, 1984; Paris & Oka, 1986). Onderzoek bij een groot aantal leerlingen uit groep 5 en 7 wees uit dat de kinderen die het curriculum volgden, in vergelijking met een controlegroep, een grotere kennis van strategieën kregen. Tevens waren ze meer vooruitgegaan op een 'cloze-test' en een 'error-detection' taak. Er was echter geen vooruitgang te constateren op een gestandaardiseerde test voor begrijpend lezen. Evenals Duffy et. al. (1987) wijten ze het ontbreken van transfer op een dergelijke toets aan het verschil tussen de concepten die ten grondslag liggen aan de traditionele testen voor begrijpend lezen en de concepten die het uitgangspunt zijn voor een strategiegerichte aanpak van begrijpend lezen.

Een instructiemethode die eveneens effectief bleek, maar in tegenstelling tot bovenstaande benaderingen bovendien wel een transfereffect opleverde, is de 'reciprocal teaching' aanpak van Brown en Palincsar (Brown & Palincsar, 1989; Palincsar en Brown, 1989). Binnen deze benadering vindt instructie plaats in de vorm van een dialoog tussen de leerkracht en de leerling over de tekst. Deze dialoog wordt gestructureerd aan de hand van vier belangrijke cognitieve activiteiten: vragen stellen, samenvatten, verduidelijken en voorspellen. Daarbij verschuift de verantwoordelijkheid geleidelijk van de leerkracht naar de leerling. De leerkracht heeft een belangrijke taak als 'hardop denker', maar uiteindelijk moet de leerling zelf de strategieën toepassen. Instructie vindt plaats in kleine groepen leerlingen van maximaal acht personen. Leerlingen die 'reciprocal teaching' ontvingen bleken na afloop van de training, in vergelijking met een controlegroep, beduidend beter te presteren op toetsen voor begrijpend lezen.

Ook Pennings, De Jong en Laudy (1989) pasten het principe van reciprocal teaching toe, namelijk bij instructie in het vinden van de hoofdgedachte van een tekst. De bevindingen van het onderzoek bij een kleine groep leerlingen uit het speciaal onderwijs waren positief. De experimentele groep was in vergelijking met een controlegroep duidelijk vooruit gegaan in het stellen van vragen en het vinden van de hoofdgedachte. Uitspraken over een eventueel transfereffect konden niet worden gedaan, omdat er geen andere toetsen voor begrijpend lezen werden afgenomen.

Bovenstaande instructiemethoden zijn voornamelijk experimentele (meestal kortdurende) interventies. Maar hoe staat het nu met instructie in begrijpend lezen in de dagelijkse praktijk van het onderwijs? Onderzoek in Nederland heeft uitgewezen dat nog maar weinig expliciete instructie wordt gegeven (Aarnoutse, 1990, 1991; Sliepen & Reitsma, 1991, Weterings & Aarnoutse, 1986). Daarnaast is uit een recente internationale vergelijkingsstudie gebleken dat de (begrijpend) leesprestaties van Nederlandse kinderen duidelijk achter blijven bij leerlingen uit andere hoog-ontwikkelde westerse landen (De Glopper & Otter, 1992). Het is nog niet duidelijk hoe deze achterstand kan worden verklaard. Toch is het niet onmogelijk dat het gebrek aan expliciete instructie in begrijpend lezen hierbij een zekere rol speelt. Uit eerder genoemde studies naar de praktijk van het onderwijs in begrijpend lezen blijkt immers dat er zeker nog wat te verbeteren is.
Een veel gehoorde vraag vanuit de onderwijspraktijk is die naar expliciete richtlijnen voor instructie in begrijpend lezen aan zwakke lezers. Geïnitieerd door deze vraag is het hier te beschrijven onderzoek opgezet. De doelstelling is te bepalen in hoeverre instructie in strategieën effectief is bij leerlingen die problemen hebben met begrijpend lezen. Vanuit de gedachte dat zwakke lezers vooral gekenmerkt worden door een passief leesgedrag, kan worden aangenomen dat zij baat kunnen hebben bij een dergelijke instructiemethode. Deze suggestie wordt tevens ondersteund door de bevindingen van eerder genoemde interventiestudies.

Een eerste vraag in dit onderzoek is of het mogelijk is om zwakke begrijpende lezers strategieën voor begrijpend lezen aan te leren. Vervolgens is het de vraag of een dergelijke instructie een positief effect heeft op de prestaties in begrijpend lezen. Om antwoord op deze vragen te krijgen is een experimenteel programma ontwikkeld voor het aanleren van strategieën aan zwakke lezers, dat vervolgens is getoetst in een interventiestudie. De verwachting was dat

door het toenemen van kennis van strategieën het begrijpend leesproces positief zou worden beïnvloed, wat weer tot uiting zou kunnen komen in betere prestaties op toetsen voor tekstbegrip.

Binnen dit onderzoek naar de effectiviteit van instructie in strategieën is nader gekeken naar één specifieke strategie: het activeren van voorkennis. Het gebruik van reeds aanwezige voorinformatie wordt binnen de recente opvattingen voor begrijpend lezen namelijk van groot belang geacht. Achtergrondinformatie is belangrijk voor het goed begrijpen van een tekst (Alverman, Smith & Readence, 1985; Schallert, 1991). De meeste instructiemethoden waarin voorkennis wordt geactiveerd zijn leerkracht-afhankelijk en tekstgericht. Dat wil zeggen dat de leerkracht instructie geeft die gericht is op het bevorderen van het begrip van één specifieke tekst, bijvoorbeeld door het geven van een 'preview', uitleg van belangrijke woorden en conceptuele schema's (Dole, Valencia, Greer & Wardrop, 1991; Graves, Cooke & LaBerge, 1983; Langer, 1984). Dergelijke tekstgerichte instructie blijkt effectief voor het begrip van een specifieke tekst. Er wordt echter geen transfer aangetoond naar een beter begrip van andere teksten.

Het is tot op heden onduidelijk in hoeverre het mogelijk is, om binnen een strategiegerichte instructiemethode, leerlingen tevens te trainen zelfstandig hun voorkennis te activeren. Daarom heeft dit onderzoek tevens als doel na te gaan in hoeverre het effectief is om zwakke lezers te leren hun voorkennis te gebruiken. De vraag is of binnen het experimentele programma het activeren van voorkennis een meerwaarde effect heeft. De verwachtingen hierover zijn niet eenduidig. Enerzijds kan er een positief effect worden verwacht, daar voorkennis van groot belang is bij tekstbegrip. Anderzijds kan het zelfstandig door de leerling activeren van voorkennis ook negatieve invloed hebben. Zo kan bijvoorbeeld het activeren van niet relevante of onjuiste voorkennis de aandacht van de inhoud van de tekst afleiden of zelfs een onjuist begrip tot gevolg hebben (Dole et al., 1991; McCormick, 1992; Schallert, 1991).

1 Methode

1.1 Procedure

Het interventie-onderzoek is uitgevoerd volgens een pretest-posttest control groep design. De periode tussen voor- en nameting bedroeg ongeveer zeven weken. Tevens zijn er vier weken na de nameting retentiemetingen uitgevoerd. Vanwege beperkingen in tijd zijn daarbij niet alle toetsen afgenomen. De toetsen zijn voornamelijk individueel of in kleine groepjes afgenomen door de onderzoeker en twee assistenten.

De leerlingen zijn verdeeld over twee groepen die zoveel mogelijk overeenkomen op relevante variabelen. Vervolgens zijn deze groepen willekeurig toegewezen aan een experimentele en een controle conditie. Binnen de experimentele groep zijn voorts twee condities onderscheiden, waaraan leerlingen willekeurig werden toegewezen. In beide condities werd de experimentele instructie gevolgd, echter slechts in één experimentele groep werd het activeren van voorkennis expliciet als strategie aangeleerd.

In de periode tussen de voor- en nameting volgde de experimentele groep 13 à 14 lessen van 30 tot 40 minuten. De lessen werden twee maal per week gegeven door de onderzoeker. Dit vond plaats buiten de klassesituatie, echter zoveel mogelijk tijdens de reguliere leeslessen. De leerlingen in de controlegroep volgden het reguliere programma in de klas. Tevens was er één (controle) bijeenkomst met de onderzoeker, waarbij in een kleine groep een tekst gelezen werd en inhoudsvragen werden beantwoord.

1.2 De onderzoeksgroep

Aan het onderzoek hebben 24 kinderen (18 jongens, 6 meisjes) deelgenomen, afkomstig van twee scholen voor speciaal onderwijs, verbonden aan een instituut voor intensieve jeugdhulpverlening. De kinderen waren door leerkrachten geselecteerd als kinderen met ernstige problemen met begrijpend lezen. Kort samengevat waren de criteria waarop de leerlingen geselecteerd werden: minimaal één jaar achterstand op het gebied van begrijpend lezen, met een minimumniveau vergelijkbaar met een groep 5-niveau (gemeten met CITO-toetsen voor begrijpend lezen, CITO, 1982); een tech-

nische leesvaardigheid van minimaal AVI-niveau 5, geen bovengrens; leeftijd 10 tot en met 12 jaar; een verbaal IQ van minimaal 80, geen bovengrens; en voldoende beheersing van de Nederlandse taal.

Zowel de experimentele als de controlegroep bestond uit 12 kinderen. De groepen zijn vergelijkbaar voor wat betreft leeftijd, verbale intelligentie (WISC-RN), woordenschat (Taaltest Voor Kinderen, maximumscore 40) en technisch lezen (Differentiële Zinnenlees Test, aantal woorden in drie minuten). Voor de controle en experimentele groep is de gemiddelde leeftijd respectievelijk 10;11 (jaar; maanden) (standaarddeviatie, *sd*=0.8) en 11;1 (*sd*=1.5); het gemiddeld verbaal IQ is respectievelijk 89.5 (*sd*=11.2) en 85.6 (*sd*=7.7); de gemiddelde score op de woordenschattoets bedraagt respectievelijk 35.0 (*sd*=3.9) en 35.9 (*sd*=2.0); en de gemiddelde score op de toets voor technisch lezen is respectievelijk 209.1 (*sd*=20.0) en 209.3 (*sd*=23.7). Er zijn geen significante verschillen tussen de gemiddelden van de controle en de experimentele groep.

Ook tussen de twee experimentele groepen zijn geen significante verschillen gevonden. Voor conditie 1 (met activatie van voorkennis) en conditie 2 (zonder activatie van voorkennis) is de gemiddelde leeftijd respectievelijk 11;2 (*sd*=2.0) en 11;1 (*sd*=0.7); het gemiddelde verbaal IQ is respectievelijk 83.0 (*sd*=9.3) en 89.5 (*sd*=5.4); de gemiddelde score op de woordenschattoets bedraagt respectievelijk 36.0 (*sd*=1.6) en 35.8 (*sd*=2.50); en de gemiddelde score op de toets voor technisch lezen is respectievelijk 219.5 (*sd*=16.9) en 199.2 (*sd*=26.5).

1.3 Meetinstrumenten

Vragenlijst kennis van strategieën. Voor het bepalen van de kennis van strategieën voor begrijpend lezen is een instrument ontwikkeld. Het betreft een vragenlijst met daarin vijftien uitspraken over het gebruik van strategieën vóór, tijdens en na het lezen. Een voorbeeld van een item is: "Voordat ik begin met lezen bedenk ik eerst waarom ik ga lezen". De leerling geeft aan of een uitspraak wel of niet op hem van toepassing is bij het lezen van informatieve teksten (ja/nee items). Van de vijftien vragen hebben twaalf uitspraken betrekking op relevante strategieën en zijn er drie niet relevant. Ieder juist relevant antwoord levert één punt op, zodat het totaal te behalen aantal punten 12 bedraagt. De score op de vragenlijst geeft een indicatie van de mate waarin een leerling zich bewust is van strategieën voor begrijpend lezen. Er is geen tijdslimiet. Bij de verschillende metingen is dezelfde vragenlijst afgenomen. De coëfficiënt voor item-consistentie (KR20) is in eerder onderzoek vastgesteld op .72 (Walraven & Reitsma, 1991).

Hoofdgedachtetest. Deze test, met parallelvormen A en B, is bedoeld voor groep 6 en 7 en bestaat uit 29 afzonderlijke teksten waarbij de hoofdgedachte moet worden gekozen uit vier alternatieven (Aarnoutse, 1984). De test wordt in twee delen afgenomen. De test beoogt een bepaald aspect van begrijpend lezen te meten, namelijk het afleiden van de hoofdgedachte uit een informatieve tekst. De maximaal te behalen score bedraagt 29. De tijdslimiet per deel is 35 minuten (totaal 70 minuten). Bij de voor- en nameting is respectievelijk de A- en B- versie afgenomen. Deze test is niet bij de retentiemeting afgenomen.

Invultest. Deze cloze-test, met parallelvormen A en B, is ontwikkeld voor groep 6. De test heeft tot doel begrijpend lezen te meten. De veronderstelling is dat bij deze test vooral inferentiële processen een belangrijke rol spelen (Aarnoutse & Buitenhuis, 1984). De test bestaat uit een informatieve tekst van ongeveer 400 woorden, waaruit na een inleidende passage elk zevende woord is weggelaten. De leerlingen moeten op de vijftig open plaatsen in de tekst een woord invullen. De score wordt bepaald door het totaal aantal goed ingevulde woorden, en bedraagt maximaal 50. De tijdslimiet is 35 minuten. Bij de voor- na- en retentiemeting is respectievelijk de A-, B- en A-versie afgenomen.

1.4 Het instructieprogramma

Het doel van het instructieprogramma is het aanleren van strategieën voor begrijpend lezen aan zwakke lezers. De strategieën doen een beroep op metacognitieve vaardigheden van de leerling: denken over hun leesproces en hier meer greep op krijgen. Kortom, ze leren hun leesgedrag te plannen, te bewaken en te evalueren. De volgende strategieën zijn daartoe in het programma opgenomen: het opstellen van een leesdoel; voorspellen op basis van tekstkenmerken; activeren van voorkennis; controleren

van leesbegrip; aandacht richten op belangrijke zaken; samenvatten en evalueren van de inhoud (Palincsar & Brown, 1989; Paris et al., 1984).

Deze strategieën worden in eenvoudige taal aan de kinderen aangeboden als zogenaamde *denkstappen*. De volgende denkstappen komen aan de orde: 1) Waarom ga ik lezen?; 2) Waar zal het over gaan?; 3) Wat weet ik er al van?; 4) Stop! Waar gaat het over?; 5) Wat is er belangrijk?; 6) Wat ben ik te weten gekomen? Deze denkstappen worden geoefend aan de hand van zakelijke informatieve teksten. De teksten worden hardop gelezen, waarbij voor, tijdens en na het lezen een dialoog over de tekst plaatsvindt aan de hand van de denkstappen.

De strategieën worden niet in één keer aangeboden, maar stap voor stap in de loop van het programma. Zo wordt in de eerste les alleen denkstap 1 behandeld. Wat later in het programma wordt stap 2 geïntroduceerd, zodat er twee strategieën worden geoefend. Enkele lessen later wordt de derde strategie aangeboden: de leerlingen moeten nu drie strategieën toepassen bij het lezen van een tekst. Het instructieprogramma heeft dus een cumulatief karakter: reeds aangeboden strategieën worden herhaald en eventueel aangevuld met een nieuwe strategie.

De denkstappen zijn, voorzien van pictogrammen, tevens weergegeven op *hulpkaarten*. Deze hulpkaarten mogen tijdens het lezen worden gebruikt door de leerlingen en bieden op die manier structuur en visuele ondersteuning bij het oefenen van de strategieën. Er zijn meerdere hulpkaarten, die elkaar vervangen. Zo staat op de eerste hulpkaart alleen denkstap 1. Deze hulpkaart wordt wat later in het programma vervangen door een tweede hulpkaart, waarop de denkstappen 1 en 2 staan. De hulpkaarten weerspiegelen zodoende het cumulatieve karakter van het instructieprogramma. In Tabel 1 is een overzicht gegeven van de strategieën en de bijbehorende denkstappen voor de leerlingen. Tevens is beknopt aangegeven aan de hand van welke activiteiten de strategieën kunnen worden geoefend.

Bij het aanbieden van de strategieën zijn principes van directe instructies gevolgd (Pearson & Dole, 1987; Veenman, 1992). Het directe-instructiemodel bestaat globaal uit drie fasen: a) presentatie en uitleg; b) inoefening en begeleiding; en c) zelfstandige of individuele oefening (Veenman, 1992). In het ontwikkelde lesprogramma is dit instructiemodel als volgt ingevuld: 1) De leerkracht (proefleider) herhaalt de leerstof uit de vorige lessen; 2) De leerkracht legt het doel van de les uit; 3) De leerkracht doet voor hoe je een bepaalde strategie toepast (modeling); 4) De leerkracht past stra-

Tabel 1
Overzicht van strategieën, denkstappen en activiteiten

Strategie	Denkstap	Activiteiten
1. Opstellen van een leesdoel.	1. Waarom ga ik lezen?	Bij verschillende teksten vooraf een leesdoel opstellen en een aanpak bedenken.
2. Verwachtingen uitspreken aan de hand van tekstkenmerken.	2. Waar zal het over gaan?	Zowel voor als tijdens het lezen verwachtingen uitspreken aan de hand van titel, subtitel, plaatjes, eerste zin en reeds gelezen tekst.
3. Het activeren van voorkennis over onderwerp.	3. Wat weet ik er al van?	Voor het lezen bedenken wat je al weet over het onderwerp.
4. Het controleren van leesbegrip.	4. Stop! Waar gaat het over?	Tijdens het lezen controle- en reflectie momenten inbouwen. Snap je het nog? Klopte het met wat je verwachtte?
5. Aandacht richten op belangrijke zaken.	5. Wat is er belangrijk?	Tijdens het lezen aandacht richten op hoofdzaken. Waar gaat het over? Wat is er belangrijk?
6. Evalueren van inhoud van gelezen tekst.	6. Wat ben ik te weten gekomen?	Na het lezen nadenken over de inhoud. Waar ging de hele tekst over? Ben je nieuwe dingen te weten gekomen?

tegieën toe samen met de leerlingen; 5) De leerlingen passen zelfstandig strategieën toe; en 6) De leerkracht laat de leerlingen de leerstof kort herhalen.

Dit model kan worden aangeduid als een directe-instructiemodel voor strategisch handelen, een model met de nadruk op processen in plaats van eindprodukten (Veenman, 1992). Binnen een dergelijk instructiemodel is er sprake van een geleidelijke overdracht van verantwoordelijkheid voor de leertaak van de leerkracht naar de leerling. Elke fase in het instructieproces vraagt om verschillende graden van verantwoordelijkheid. Zo ligt bij de presentatie en uitleg de verantwoordelijkheid vooral bij de leerkracht. Bij het aanbieden van een nieuwe strategie bijvoorbeeld, doet de leerkracht voor hoe de strategie toegepast kan worden, en verwoordt daarbij hardop de denk- en oplossingsprocessen. Wanneer strategieën echter al vaker zijn geoefend, worden de leerlingen steeds meer gestimuleerd de strategieën zelfstandig toe te passen.

In het ontwikkelde instructieprogramma is de actieve participatie van de leerling tijdens de fase van inoefening extra gestimuleerd door principes van 'reciprocal teaching' (wederkerig onderwijzen) te volgen (Brown & Palincsar, 1989; Palincsar & Brown, 1989). De essentie van deze aanpak is een dialoog over de tekst waarbij de verantwoordelijkheid geleidelijk verschuift van de leerkracht naar de leerling. In het ontwikkelde programma nemen de leerlingen afwisselend de rol van leerkracht over, doen de strategieën voor en vragen de andere leerlingen de strategieën toe te passen bij het lezen.

De gemanipuleerde strategie 'activatie van voorkennis' is slechts bij één experimentele conditie aangeleerd. Aan de hand van de vraag "Wat weet ik er al van?", werden de leerlingen gestimuleerd om, voordat ze gingen lezen, eerst goed na te denken en op te schrijven wat ze al wisten over het onderwerp. Bij het voordoen van deze strategie door de leerkracht werd sterk benadrukt dat alleen relevante informatie moest worden opgeroepen en gebruikt. Zwakke lezers neigen er namelijk toe om alles op te noemen wat in hun hoofd opkomt, en kunnen daarbij moeilijk hoofd- en bijzaken van elkaar onderscheiden (McCormick, 1992).

2 Resultaten

De prestaties van de controlegroep en de experimentele groep (als geheel) op de verschillende meetmomenten zijn weergegeven in Tabel 2.

Allereerst is met behulp van *t*-toetsen gecontroleerd voor eventuele verschillen op de voortoetsen tussen de experimentele en de controlegroep. Daarbij zijn geen significante verschillen geconstateerd, zodat mag worden verondersteld dat de groepen niet verschillen bij aanvang van de interventie.

Uit een vergelijking van de resultaten op de voor- en natoetsen blijkt dat de experimentele groep op alle toetsen meer vooruit is gegaan dan de controlegroep. Deze tendens is getoetst met behulp van een MANOVA (drie afhankelijke variabelen) met twee herhaalde metingen. Deze analyse liet een significant interactie-effect zien (groep x herhaalde meting), Wilks' Lambda $F(3,20)=12.9$, $p<.001$. Hieruit kan worden geconcludeerd dat de instructie in stra-

Tabel 2
Gemiddelde scores (M) en standaardafwijkingen (SD) van de controle- en de experimentele groep

	Controlegroep n=12			Experimentele groep n=12		
	Voor	Na	Retentie	Voor	Na	Retentie
Vragenlijst						
M	5.4	5.3	4.5	6.9	11.2	9.3
SD	3.6	3.3	3.2	2.9	1.2	1.6
Hoofdgedachte						
M	11.4	11.6	-	8.0	13.3	-
SD	7.0	8.1	-	3.2	3.3	-
Invultest						
M	23.7	20.4	25.3	22.7	27.3	30.0
SD	6.1	4.8	6.8	6.0	5.9	5.5

tegieën een positief effect heeft gehad op de prestatie in begrijpend lezen. Het interactie-effect is tevens nagegaan voor de afzonderlijke afhankelijke variabelen. Er was sprake van significante interactie-effecten op respectievelijk de Vragenlijst voor kennis van strategieën, $F(1,22)= 19.5$, $p<.001$; de Invultest, $F(1,22)= 21.2$, $p<.001$; en de Hoofdgedachtetest, $F(1,22)= 6.7$, $p<.05$. Teneinde te bepalen hoeveel variantie kan worden verklaard door het effect van interventie, zijn tevens de 'estimated treatment effects' ω (omega) berekend. Voor de Vragenlijst, Invultest en Hoofdgedachtetest waren dat respectievelijk, ω=.43, ω=.37 en ω=.12.

Om te bepalen of deze positieve effecten van de interventie ook vier weken na de nameting nog aanwezig waren, is een MANOVA (met twee afhankelijke variabelen) met drie herhaalde metingen uitgevoerd. Uit deze analyse bleek dat de effecten van de interventie nog steeds konden worden aangetoond. Er werd namelijk een significant interactie-effect (groep x herhaalde meting) gevonden, Wilks' Lambda $F(4,86)= 7.8$, $p<.001$. Ook univariate variantie-analyses met drie herhaalde metingen wezen op significante interactie-effecten, respectievelijk voor de Vragenlijst, $F(2,22)= 14.0$, $p<.001$; en de Invultest, $F(2,22)= 8.1$, $p<.001$. Geschatte interventie-effecten bij de retentiemeting waren voor de Vragenlijst ω=.39 en voor de Invultest ω=.12.

Vervolgens is nagegaan of de twee experimentele groepen in verschillende mate van de instructie geprofiteerd hebben. De prestaties van de twee condities op de verschillende meetmomenten zijn weergegeven in Tabel 3.

Allereerst is met behulp van *t*-toetsen vastgesteld dat er geen significante verschillen waren tussen de twee experimentele groepen bij aanvang van de interventie. Uit Tabel 3 is verder op te maken dat beide experimentele groepen een vooruitgang laten zien tussen de voor- en nameting. In een MANOVA met herhaalde metingen werd dit bevestigd: er is sprake van een significant hoofdeffect voor herhaald meten, $F(3,8)= 19.8$, $p<.001$. De gegevens laten een klein verschil in vooruitgang zien tussen beide groepen; de vooruitgang van de experimentele groep 'met activatie voorkennis' is iets groter. In de MANOVA is deze tendens enigszins terug te zien in de grootte van het interactie-effect (groep x herhaald meten), Wilks' Lambda $F(3,8)= 2.6$, $p=.12$. Een duidelijk differentieel effect tussen strategie-instructie met of zonder activatie van voorkennis kan echter niet worden aangetoond.

Vervolgens is nagegaan of er op langere termijn verschillen te zien zijn in het effect van de verschillende vormen van instructie. Bij beide experimentele groepen is ook enkele weken na de interventie nog een positief effect van de instructieperiode waar te nemen. Een MANOVA met drie herhaalde metingen levert een significant hoofdeffect voor herhaalde meting op, Wilks' Lambda $F(4,38)$, $p< .001$. Er zijn echter geen verschillen in effecten tussen de groepen te constateren. Ook hier is echter de *F*-waarde van het interactie-effect redelijk te noemen, Wilks' Lambda $F(4,38)= 2.3$, $p=.08$. Met behulp van een MANOVA is het verschil tussen de groepen op de Vragenlijst voor strategieën nog aan nadere toetsing onderworpen. De 'voorkennis-groep' had na de interventie meer kennis van strategieën, vergeleken met de andere experimentele groep. Er was sprake van

Tabel 3
Gemiddelde scores (M) en standaardafwijkingen (SD) van de twee experimentele groepen

	Met activatie voorkennis *n*=6			Zonder activatie voorkennis *n*=6		
	Voor	Na	Retentie	Voor	Na	Retentie
Vragenlijst						
M	6.0	11.7	10.0	7.8	10.7	8.7
SD	3.5	0.5	1.6	2.1	1.5	1.5
Hoofdgedachte						
M	8.0	14.3	-	8.0	12.3	-
SD	3.9	3.4	-	2.8	3.2	-
Invultest						
M	21.3	28.3	30.3	24.0	26.3	29.7
SD	2.2	3.9	5.2	8.3	7.7	6.4

een significant interactie-effect (groep x herhaalde meting), $F(2,10)= 4.7$, $p<.05$. De hoeveelheid door interventie-effect verklaarde variantie op deze vragenlijst is $\omega=.43$.

3 Discussie

Is het mogelijk om zwakke lezers strategieën voor begrijpend lezen aan te leren? Op grond van de resultaten uit het hier beschreven onderzoek mag deze vraag positief worden beantwoord. De leerlingen die expliciete instructie kregen in het gebruik van strategieën, bleken na afloop van de interventie meer kennis van strategieën te hebben gekregen dan de controlegroep. Tevens bleek de toename redelijk stabiel. Deze bevindingen zijn in overeenstemming met de conclusies uit andere interventiestudies waarin metacognitieve strategieën werden aangeleerd (Paris et al., 1984; Duffy et al., 1987).

Vooruitgang in kennis van strategieën is echter geen doel op zich. Waar het vanzelfsprekend om gaat is of instructie in strategieën een positieve uitwerking heeft op de prestaties in begrijpend lezen. De achterliggende theoretische veronderstelling hierbij is, dat vermeerdering van strategiekennis en 'reading awareness' een positieve invloed heeft op het proces van begrijpend lezen. Uit de resultaten van dit onderzoek is af te leiden dat de kinderen die expliciete instructie ontvingen beter zijn gaan presteren op toetsen voor begrijpend lezen dan de kinderen die het reguliere leesonderwijs hebben gevolgd. Dit effect van de instructie bleek vier weken na de interventie nog aanwezig. Of deze vooruitgang in tekstbegrip direct te herleiden is tot de vermeerdering van strategiekennis en -gebruik, is een vraag die aan de hand van dit onderzoek niet kan worden beantwoord. Het is echter zeer plausibel te veronderstellen dat de directe instructie in strategieën de leerlingen tot meer actieve en bewuste lezers heeft gemaakt. Deze verandering in attitude kan een positief effect hebben gehad op de leesprestaties.

Deze verbetering van de prestaties mag worden geïnterpreteerd als een duidelijk transfereffect. De gebruikte toetsen voor begrijpend lezen deden immers geen expliciet beroep op het gebruik van strategieën, zodat grote vooruitgang op deze toetsen niet zonder meer te verwachten was.

Een dergelijk effect werd ook gevonden in andere studies waarin het directe instructiemodel voor strategisch handelen werd gevolgd, zoals het 'reciprocal teaching'-onderzoek van Palincsar en Brown (1989). Zou deze specifieke didactische aanpak wellicht een grote rol spelen in het uiteindelijk effect van de interventie? Mogelijk worden de leerlingen door directe instructie en in het bijzonder door reciprocal teaching gestimuleerd om actief bij de tekst betrokken te zijn, hardop te denken en probleemoplossend bezig te zijn. Dit bevordert mogelijk de ontwikkeling van metacognitieve vaardigheden (Veenman, 1992). Over de effecten van respectievelijk de instructievorm, de inhoudelijke aspecten van de instructie (specifieke strategieën) en de combinatie van beide is tot op heden nog niets bekend. Het is echter interessant te weten of het succes van een bepaalde interventie voornamelijk terug te voeren is op de didactische aanbieding of juist op de inhoud: de specifieke vaardigheden en strategieën.

Een factor die eventueel ook zou kunnen meespelen bij de vooruitgang van de experimentele groep, is de extra aandacht en extra leesoefening die zij hebben ontvangen in vergelijking met de controlegroep. Deze factor lijkt echter niet van groot belang, omdat in het onderwijs aan deze kinderen al veel individueel gericht onderwijs wordt gegeven. Tevens werd de instructie zoveel mogelijk tijdens reguliere leeslessen gegeven, waarin veel aandacht wordt besteed aan het (technisch) lezen.

In dit onderzoek zijn zowel strategie-instructie zonder, als strategie-instructie inclusief activatie van voorkennis effectief gebleken. De vooruitgang in strategiekennis na afloop van de interventie is vergelijkbaar voor beide groepen. Ook op de toetsen voor begrijpend lezen vertoonden beide experimentele groepen een vooruitgang. Een eventueel meerwaarde-effect van voorkennisactivatie is in dit onderzoek niet aangetoond.

Bij deze laatste bevinding moet allereerst worden opgemerkt dat het aantal proefpersonen in de experimentele condities dermate klein was, dat interindividuele verschillen een sterke invloed kunnen hebben gehad. Wellicht was deze invloed sterker dan de systematische

effecten van de experimentele condities. Door het beperkte aantal proefpersonen is de power, de kans om eventueel aanwezige verschillen daadwerkelijk aan te tonen, klein. Om deze power te vergroten zou een groter aantal proefpersonen per conditie gewenst zijn (Stevens, 1992).

Een andere kanttekening is dat de gebruikte afhankelijke variabelen niet expliciet vragen om het activeren van voorkennis; de toetsen zijn dientengevolge niet erg gevoelig voor eventuele veranderingen in gebruik van voorkennis. Het afnemen van toetsen die een duidelijkere relatie hebben met de gemanipuleerde instructievorm zou ook de power kunnen vergroten (Stevens, 1992).

In de inleiding is reeds gesuggereerd dat het activeren van voorkennis ook negatieve invloed zou kunnen hebben op het leesgedrag. Het efficiënt gebruik maken van voorkennis is geen makkelijke opgave voor zwakke lezers. Zo is het activeren van voorkennis een tijdconsumerende bezigheid, die daardoor een zwaar beroep doet op de aandachts- en geheugencapaciteit. Tevens is het voor deze zwakke lezers moeilijk om beslissingen te nemen over het al dan niet relevant zijn van informatie voor een specifieke tekst. De gegevens uit dit onderzoek geven echter geen aanleiding te veronderstellen dat het activeren van voorkennis een negatieve invloed heeft gehad op de prestaties. De vooruitgang van de 'voorkennis-groep' is immers niet achtergebleven bij de andere groep. Het is zeer wel mogelijk dat bij een langere interventieperiode de eventuele meerwaardeeffecten meer zichtbaar zouden worden. Het zelfstandig en efficiënt leren activeren van voorkennis vraagt wellicht meer tijdsinvestering dan in dit onderzoek beschikbaar was.

Tot slot zou een mogelijke verklaring voor het ontbreken van een voorkennisactivatie effect kunnen liggen in de instructievorm. Zoals eerder vermeld lag de nadruk op het zelfstandig toepassen van strategieën door de leerling. Activatie van achtergrondinformatie is echter alleen dan nuttig wanneer het relevante informatie betreft en op een gestructureerde manier gebeurt. Wellicht kunnen tien- tot twaalfjarige zwakke lezers de verantwoordelijkheid voor een dergelijke taak nog niet aan. Mogelijk moet bij dit deel van het leesproces de verantwoordelijkheid vooral bij de leerkracht (blijven) liggen.

Wat is de relevantie van dit onderzoek voor de praktijk van het onderwijs? Ten eerste heeft dit onderzoek aangetoond dat het zinvol is om zwakke lezers te leren hun leesgedrag beter te sturen, te bewaken en te evalueren. Door strategie-instructie krijgen leerlingen meer inzicht in *hoe* ze verschillende teksten kunnen aanpakken. Tevens heeft instructie in strategieën een positieve invloed op de prestaties.

Voor zwakke lezers lijkt het van belang dat deze strategieën stapsgewijs en expliciet via directe instructie worden aangeboden. Het principe van 'reciprocal teaching' is daarvoor een geschikte didactische aanbiedingsvorm, omdat het leerlingen dwingt over hun leesgedrag na te denken en actief de tekst te verwerken. Het hardop denken en verwoorden van aanpakgedrag speelt daarbij een belangrijke rol. Bovendien is het bevorderlijk voor het zelfvertrouwen van de leerling om te zien dat ook anderen moeilijkheden kunnen hebben bij het lezen en begrijpen, en om te zien wat anderen doen om de taak zo goed mogelijk uit te voeren. De leerlingen kunnen elkaar op die manier aanvullen en ondersteunen.

Strategieën voor begrijpend lezen kunnen natuurlijk ook in grotere groepen of klassikaal worden onderwezen, zeker wanneer de leesniveaus niet al te ver uit elkaar liggen. Daarbij moet echter wel rekening worden gehouden met het feit dat de actieve (effectieve?) leertijd van de individuele leerling afneemt. Een langere periode van instructie is dan belangrijk. Het is zelfs aan te raden om instructie in strategisch en zelfregulerend leesgedrag structureel in te bouwen in het leesonderwijs. Daarbij is het overigens wenselijk om instructie in strategieën voor begrijpend lezen niet te isoleren als een 'apart vak'. Juist door integratie in allerlei vakken waarbij informatie moet worden verwerkt, zoals bijvoorbeeld wereldoriëntatie, wordt instructie nuttig en effectief voor de leerlingen. Goed begrijpend lezen is dan ook functioneel.

Literatuur

Aarnoutse, C.A.J. (1984). *Hoofdgedachtetest; test voor begrijpend lezen bestemd voor het vierde en vijfde leerjaar van het basisonderwijs. Verantwoording en handleiding.* Nijmegen: Berkhout.

Aarnoutse, C. A. J. (1990). Begrijpend lezen: hoe onderwijs je dat? In C. A. J. Aarnoutse & M. J. M. Voeten (red.), *Gaat en onderwijst*. Liber amicorum voor Dr. M. J. C. Mommers. Tilburg: Zwijsen.

Aarnoutse, C. A. J. (1991). Begrijpend lezen in het basisonderwijs. In P. Reitsma & A. M. A. Walraven (Eds.), *Instructie in begrijpend lezen*. RAIN-1. Delft: Eburon.

Aarnoutse, C. A. J., & Buitenhuis, A. F. (1984). *Invultest; test voor begrijpend lezen bestemd voor het vierde leerjaar van het basisonderwijs. Verantwoording en handleiding*. Nijmegen: Berkhout.

Alverman, D. E., Smith, L. C., & Readence, J. E. (1985). Prior knowledge activation and the comprehension of compatible and incompatible text. *Reading Research Quarterly, 20*(4), 420-436.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson (ed.), *Handbook of reading research*. New York: Longman.

Bristow, P. S. (1985). Are poor readers passive readers? Some evidence, possible explanations, and potential solutions. *The Reading Teacher, 39*, 3, 318-325.

Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., & Palincsar, A. S. (1989). Guided, coöperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning and instruction. Essays in honor of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum Associates.

CITO (1982), *Toetsen begrijpend lezen, leerjaar 3, 4 en 5*. Arnhem: Centraal Instituut voor Toets Ontwikkeling.

Dole, J. A., Valencia, S. W., Greer, E. A., & Wardrop, J. L. (1991). Effects of two types of prereading instruction on the comprehension of narrative and expository text. *Reading Research Quarterly, 26*(2), 142-159.

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M., Vavrus, L., Wesselman, R., Putnam, J., & Bassari, D. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly, 22*, 347-368.

Forrest-Pressley, D. L., & Gillies, L. A. (1983). Children's flexible use of strategies during reading. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research, educational applications*. New York: Springer-Verlag.

Glopper, K. de, & Otter, M. E. (1992). Leesprestaties Nederlandse kinderen vallen tegen. *Didaktief, 22*(10), 4-6.

Graves, M. F., Cooke, C. L., & LaBerge, M. J. (1983). Effects of previewing short stories on low ability junior high school students' comprehension, recall and attitudes. *Reading Research Quarterly, 18*(3), 262-276.

Langer, J. A. (1984). Examining background knowledge and text comprehension. *Reading Research Quarterly, 19*(4), 468-481.

McCormick, S. (1992). Disabled readers' erroneous responses to inferential comprehension questions: description and analysis. *Reading Research Quarterly, 27*(1), 54-77.

Palincsar, A. S., & Brown, A. L. (1989). Instruction for self-regulated reading. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the Thinking Curriculum: Current Cognitive Research*. 1989 Yearbook of the Association for Supervision and Curriculum Development.

Paris, S. G., & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review, 6*, 25-56.

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*, 6, 1239-1252.

Pearson, P. D., & Dole, J. A. (1987). Explicit comprehension instruction. A review of research and a new conceptualization of instruction. *The Elementary School Journal, 88*, 2, 151-165.

Pennings, A., Jong, T. de, & Laudy, M. (1989). Het bevorderen van het gebruik van leesstrategieën via een dialogische interventieprocedure bij kinderen met leerproblemen. *Tijdschrift voor Orthopedagogiek, 28*, 541-551.

Schallert, D. .L. (1991). Perspective on recent research on reading comprehension instruction. In P. Reitsma & A. M. A. Walraven (Eds.), *Instructie in begrijpend lezen*. RAIN-1. Delft: Eburon.

Sliepen, S., & Reitsma, P. (1991). Instructie in begrijpend lezen in LOM-scholen? In B. H. A. M. van Hout-Wolters & L. F. W. de Klerk (Eds.), *Onderwijsleerprocessen; cognitie en motivatie*. Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de UvA.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences*. Second edition. Hillsdale, NJ: LEA.

Veenman, S. (1992). Effectieve instructie volgens het directe instructiemodel. *Pedagogische Studiën, 69*, 242-269.

Walraven A. M. A., & Reitsma, P. (1991). Tekstbegrip en metacognitie bij jonge en zwakke lezers. In B. H. A. M. van Hout-Wolters & L. F. W. de Klerk (Eds.), *Onderwijsleerprocessen: cognitie en motivatie*. Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de UvA.

Weterings, A. C. E. M., & Aarnoutse, C. A. J. (1986) De praktijk van het onderwijs in begrijpend lezen. *Pedagogische Studiën, 63*, 387-400.

Manuscript aanvaard 21-4-1993

Auteurs

A. M. A. Walraven, P. Reitsma en **E. J. Kappers** zijn respectievelijk als orthopedagoge, psycholoog en orthopedagoog verbonden aan het Paedologisch Instituut te Duivendrecht.

Adres: Paedologisch Instituut, Postbus 303, 1115 ZG Duivendrecht.

Abstract

Instruction in strategies for reading comprehension to poor readers

A. M. A. Walraven, P. Reitsma & E. J. Kappers. Pedagogische Studiën, 1993, 70, 298-308.

The present study examined the effects of teaching comprehension-fostering strategies to pupils with severe problems in reading comprehension. Moreover, the possible surplus-effect of activating background knowledge was investigated. Subjects were 24 Dutch children from schools for children with learning disabilities, aged 10-12 years. A pretest-posttest control group design was used. Twelve children, in groups of three, received a strategy training. Six children were taught to also activate their background knowledge before reading, whereas six others were not. Principles of direct instruction and reciprocal teaching were employed. Twelve children, following regular reading lessons, served as a controle group. The results indicate that children who received strategy training improve their awareness of strategies and also improve their performances on a cloze-test and a main-idea test. No additional effect of activating background knowledge was found.

Kroniek

Voorschoolse stimuleringsprojecten

De 'Kroniek' is als rubriek van Pedagogische Studiën onder meer bedoeld om een overzicht te schetsen van de ontwikkelingen die zich op een bepaald pedagogisch of onderwijskundig terrein voordoen, waarbij deze ontwikkelingen niet per se of uitsluitend de wetenschappelijke ontwikkelingen op dat terrein hoeven te zijn. Een terrein waarop zich interessante ontwikkelingen voordoen is dat van de voorschoolse stimuleringsprojecten voor kinderen uit sociaal-economisch zwakke milieus (vooral allochtone kinderen). Een overzicht van deze ontwikkelingen mag in een nummer waarin (de evaluatie van) het onderwijsvoorrangsbeleid centraal staat niet ontbreken.

De voorschoolse periode als een tijdsspanne waarin activiteiten ontplooid moeten worden ter stimulering van de ontwikkeling van kinderen uit sociaal-economisch zwakke milieus, in het bijzonder uit allochtone milieus, komt steeds meer in de belangstelling. Het Ministerie van WVC vervult daarbij een voortrekkersrol; initiatieven voor het ontwikkelen en verspreiden van voor allochtone kinderen bedoelde stimuleringsactiviteiten worden sterk aangemoedigd èn gesubsidieerd. In het kader van het onderwijsvoorrangsbeleid bestaat eveneens belangstelling voor de voorschoolse periode: scholen en welzijnsinstellingen voeren gezamenlijk activiteiten uit bedoeld voor kinderen verkerend in deze periode. In deze kroniek wordt getracht een beeld te schetsen van de ontwikkelingen die in Nederland met betrekking tot de voorschoolse periode gaande zijn; wij doen dat aan de hand van een aantal (min of meer) recente publikaties op dit gebied.

Voorschools?

Wij moeten allereerst enige aandacht schenken aan een terminologische kwestie. Naar Nederlandse maatstaven gemeten begint de voorschoolse periode bij de geboorte van een kind en eindigt deze bij de intrede in de basisschool (in de regel op vierjarige leeftijd dus). Met andere woorden: voorschools beslaat de periode 0-4 jaar. In veel publikaties over de voorschoolse periode wordt gebruik gemaakt van Amerikaanse literatuur, en ook wel literatuur uit andere landen; dáár betekent voorschools echter de periode 3-5 jaar en dat onafhankelijk van het schooltype dat de kinderen van deze leeftijdsgroep (al dan niet) bezoeken: het is de periode van zo'n twee jaar die voorafgaat aan de zgn. Kindergarten, een voorbereidingsklas op de eerste klas lager onderwijs, die bezocht wordt door vijfjarige kinderen. In deze zin verschillen de Nederlandse en Amerikaanse definities van de voorschoolse periode dus nogal van elkaar en het is van groot belang steeds bij elk effectiviteitsonderzoek van programma's uit de voorschoolse periode na te gaan over welke groep van kinderen men nu eigenlijk spreekt. Zo laten zich de Amerikaanse preschool-programma's voor zover deze in kindercentra (veelal Head Start centra) worden uitgevoerd (leeftijd meestal 3-4-5 jaar), het best vergelijken met de Nederlandse kleuterschoolprogramma's (voor 4-5-6 jaar).

We zijn er nog niet wat de terminologie aangaat, want de laatste jaren worden ook geheel andere definities van voorschools gehanteerd: alle jaren voorafgaand aan de laagste klas (naar Nederlandse begrippen: voorafgaand aan groep 3 BaO). Bijvoorbeeld Rispens en Van der Meulen (1992) hanteren deze definities; daaronder valt dan dus óók het kleuteronderwijs (groep 1 en 2 BaO), of in Amerika de Kindergarten en de eventueel daaraan voorafgaande schooltypes (nursery schools voor drie- en vierjarige middle class kinderen) en preschool compensatory programmes (in Head Start centres bijvoorbeeld) voor kinderen uit de lagere milieus. De stimuleringsactiviteiten voor de gehele periode (0-6 jaar) worden ook wel aangeduid als 'early childhood education'. Nogmaals, deze terminologische exercitie leert ons dat het van belang is steeds aan te geven voor welke leeftijdsgroep een bepaalde activiteit of een bepaald programma bestemd is alvorens er verder iets van te zeggen. Daarmee is tevens duidelijk geworden dat we niet 'zo maar' de effecten van het ene voorschoolse

programma met die van het andere mogen vergelijken. Een zogenaamd voorschools stimuleringsprogramma (voor vier- en vijfjarigen) náást het kleuterschoolprogramma (voor vier- en vijfjarigen) zoals in Nederland het geval kan zijn, is heel wat anders dan een voorschools programma zonder kleuteronderwijs voor de vierjarigen en met kleuteronderwijs (Kindergarten) voor de vijfjarigen, zoals veelal in Amerika het geval is.

OpStap

Het belang van de stimulering van de ontwikkeling van kinderen uit sociaal-economisch lagere milieus gedurende de voorschoolse periode (0-6 jaar) wordt allerwegen van belang gevonden (Leseman, 1989), maar krijgt desondanks toch pas sedert enige jaren op grotere schaal vorm, in Nederland althans. Een belangrijke katalysator in het vormgevingsproces is het OpStap-programma: de introductie en verspreiding daarvan zijn door WVC en de Amsterdamse Averroësstichting geregeld. Het OpStap-programma is een gezinsactiveringsprogramma voor kinderen van vier tot zes jaar, gebaseerd op het Israëlische HIPPY-programma (Lombard, 1981). Het is, in de Nederlandse situatie althans, een programma dat naast het kleuterschoolprogramma (groep 1 en 2 BaO) wordt aangeboden ten behoeve van, in de eerste plaats, allochtone kinderen. Het programma is zowel in analytische zin geëvalueerd, door groepen experts (Pels, 1990; Cie OpStap-Opnieuw, 1992; Vallen, 1992) als in empirische zin (Eldering & Vedder, 1992). Bovendien is aan de moeders die met het programma werken naar hun mening over het programma gevraagd en datzelfde is gedaan bij leerkrachten die kinderen in hun klas hebben die met het programma begeleid zijn (Felix, 1992). Ten slotte is een verkennende studie verricht naar mogelijke alternatieven voor het OpStap-programma (Eldering & Choenni-Gobardhan, 1990) en is het eerste congres in Nederland over uitsluitend gezinsactivering georganiseerd tijdens hetwelk het OpStap-programma een belangrijke rol speelde (Rispens & Van der Meulen, 1992). Kortom, er is nogal wat over en rond dit programma geschreven en gesproken.

Het OpStap-programma is, zoals gezegd, een bewerking van het Israëlische HIPPY-programma; het bestaat uit zestig taken (veelal op basis van werkbladen) die gedurende twee jaar door een (allochtone) moeder met haar kind uitgevoerd moeten worden. De moeder wordt daartoe getraind (via rollenspel) door een para-professional, de zgn. buurtmoeder, en via groepsbijeenkomsten. De taken hebben betrekking op de ontwikkeling van de zintuiglijke discriminatie, de taal- en begripsontwikkeling en op de ontwikkeling van het proleemoplossend vermogen.

De uitkomsten van de analytische evaluatie van het programma zijn in hoofdlijnen positief: het programma wordt zowel qua methodiek als qua inhoud positief beoordeeld door de experts. Echter, op tal van punten worden verbeteringen voorgesteld: met de cultuur van de allochtone gezinnen moet veel meer rekening worden gehouden, de wetenschappelijke verantwoording verdient verbetering (bijvoorbeeld de keuze van de concrete doelen, van de leeftijdsopbouw e.d.), de relatie met het kleuteronderwijs moet geëxpliciteerd worden, de flexibiliteit van het programma moet vergroot worden, de taalkundige aspecten moeten aanzienlijk verbeterd worden en in relatie worden gebracht met de cognitieve ontwikkeling (Pels, 1990; Vallen, 1992). Op basis van deze en vele andere punten is een voorstel uitgewerkt voor het aanpassen van het programma (Cie OpStap-Opnieuw, 1992). Deze aanpassing vindt momenteel plaats onder leiding van de hoogleraar Rispens (Rijksuniversiteit Utrecht).

Uit bevraging van de moeders die met het OpStap-programma werken en uit bevraging van de leerkrachten bij wie de kinderen die aan het programma deelnemen in de klas zitten, komt een positief beeld van de effecten van het programma naar voren: de ontwikkeling van de kinderen wordt door het programma gestimuleerd (Felix, 1992). De positieve opinie van zowel de experts als van de moeders en de leerkrachten staat in tegenstelling tot de uitkomsten van de empirische evaluatie van het programma. In een quasi-experimentele proefopzet, aangevuld met regressie-analyse, konden Eldering en Vedder (1992) slechts zeer geringe effecten van het OpStap-programma aantonen. Van hun uitstekend opgezette evaluatie geven zij in de genoemde publikatie een

gedetailleerd verslag waaruit blijkt dat de OpStap-kinderen zich alleen een wat betere oog-hand coördinatie eigen maakten (in vergelijking met de kinderen uit de controlegroep), maar bijvoorbeeld niet hoger scoorden op een IQ-test, noch op algemene (Nederlandstalige) taalvaardigheidstoetsen. Aangetoond kon wel worden dat er een positief verband bestaat tussen de mate waarin het programma was ingevoerd en de hoogte van de effecten. De evaluatoren concluderen dat aanpassing van het programma nodig is, o.a. intensivering van de uitvoering en voorkoming van uitval en flexibilisering van het starre geheel van activiteiten. Zij wijzen er bovendien op dat hun aanbevelingen voor aanpassing beter te realiseren zijn in een centre-based dan in het home-based OpStap-programma.

Mogelijke alternatieven voor het OpStap-programma zijn onderzocht door Eldering en Choenni-Gobardhan (1990). In een verkennende studie bestudeerden zij, mede op basis van hun lange-termijneffecten, een zevental gezinsactiveringsprogramma's. Zij concludeerden dat, hoewel de effecten van deze programma's over de jaren heen afnemen (op intelligentie, taal- en cognitieve ontwikkeling) een vijftal progamma's ook na enkele jaren nog effect sorteerde, waaronder het HIPPY-programma (de Israëlische, oorspronkelijke versie van OpStap). Hun studie laat ook zien dat het dus niet per se het HIPPY-programma is dat als enige effect sorteert (althans destijds in Israël).

Inmiddels wordt het OpStap-programma op grote schaal in Nederland gebruikt, wordt de verdere verspreiding ervan door WVC en de gemeenten, soms de onderwijsvoorrangsgebieden, betaald en door de Averroèsstichting ondersteund. En dat alles in afwachting van de nieuwe versie van het programma, die over enkele jaren gereed moet zijn.

Uitstralingseffecten van OpStap

Alle zojuist geschetste ontwikkelingen met betrekking tot het OpStap-programma hebben een zeker 'uitstralingseffect' gehad op de ontwikkeling van andere voorschoolse stimuleringsprogramma's. De eerder genoemde Averroèsstichting heeft een tweetal 'echt' voorschoolse programma's in ontwikkeling genomen: OpStapje en Klimrek (Bekkers, Van Embricqs & Van Loggem, 1992).

OpStapje hanteert dezelfde methodiek als OpStap (met o.a. de buurtmoeders); het is een gezinsactiveringsprogramma dat door de moeder en het (allochtone) kind uitgevoerd moet worden. Het is bedoeld voor kinderen van twee en drie jaar. De belangrijkste componenten zijn (ook hier weer): zintuiglijke discriminatie, taal- en begripsontwikkeling.

Het Klimrek-programma is bedoeld voor kinderen van anderhalf tot drie jaar; het is gebaseerd op het Mother Child Home Programme van Levenstein (Madden, O'Hara & Levenstein, 1984), eveneens een gezinsactiveringsprogramma, waarmee de cognitieve en de sociaal-emotionele ontwikkeling gestimuleerd worden. Overigens worden ook activiteiten uit dit programma in OpStapje verwerkt.

Een ander uitstralingseffect heeft te maken met de al eerder genoemde aanbeveling van Eldering en Vedder om met centre-based programma's te werken in plaats van met programma's die in de gezinssituatie uitgevoerd moeten worden. Indachtig de goede langetermijnresultaten van het High/Scope Preschool Curriculum, ook wel het Perry Preschool Curriculum genoemd, ontwikkeld door Weikart c.s. (cf. Barnett, 1985), is door een groep deskundigen aan WVC en O&W het advies uitgebracht om dit programma om te werken voor de Nederlandse situatie en te bestemmen voor het werken met driejarige (allochtone) kinderen in peuterspeelzalen en daarop aansluitend de basisideeën van dit curriculum (het zgn. actieve leren via een aantal kernervaringen) voort te zetten in de laagste groepen van het basisonderwijs ('Allochtone Kleuters Meer In Actie', 1993). Nauwe samenwerking tussen peuterspeelzaal en basisschool is hierbij gewenst. Tevens wordt WVC en O&W geadviseerd om het kleuterschoolprogramma zoals dat door het APS is ontwikkeld onder leiding van Jansen-Vos, het programma 'Basisontwikkeling' (1990), te doen uitbouwen ten behoeve van het peuterspeelwerk en daarna eventueel de effecten van het High/Scope Curriculum en het Basisontwikkelingsprogramma met elkaar te vergelijken. Het advies baseert zich ten opzichte van 'Basisontwikkeling' niet op evalua-

tiegegevens, omdat deze niet voorhanden zijn. Het High/Scope Curriculum wordt reeds jarenlang geëvalueerd. De (zeer) lange-termijneffecten behelzen zaken als 'minder vaak verwezen naar het SO', 'vaker een diploma VO'. 'vaker tertiair onderwijs', 'minder vaak gearresteerd', 'vaker betaald werk op 19e jaar' e.d. (voor een overzicht raadplege men het al genoemde advies 'Allochtone kleuters meer in actie', 1993).

Een laatste uitstralingseffect is dat van het 'InStapje'-programma. Dit is een gezinsactiveringsprogramma, bedoeld voor (allochtone, met name Surinaamse) kinderen van nul tot twee jaar. Het werkt evenals OpStap met buurtmoeders (para-professionals), die de moeders scholen. Het programma verkeert nog in een experimenteel stadium. De Universiteit van Nijmegen en de Stichting Krosbe te Rotterdam werken eraan (Riksen-Walraven, 1992).

Overige voorschoolse projecten

In deze paragraaf bespreken wij nog kort enkele andere voorschoolse projecten (voor nul- tot zesjarige kinderen deze keer), die niet tot de uitstralingseffecten van OpStap gerekend kunnen worden.

Het programma 'Spelvoorlichting aan huis' lijkt sterk op het Mother-Child Home Programme van Levenstein (zie vorige paragraaf). Een spelleídster bezoekt eenmaal per week een gezin met jonge, allochtone kinderen en speelt met deze kinderen, daarbij ook de moeder betrekkend. Soms worden de moeders ook voor groepsbijeenkomsten uitgenodigd. Het project wordt ondersteund door de Stichting Spel- en Opvoedingsvoorlichting. Op verzoek van WVC heeft een evaluatiecommissie van deskundigen het programma tegen het licht gehouden. (Ook) deze commissie oordeelt in grote lijnen positief over het project, uiteraard met de nodige aanbevelingen voor de verbetering ervan, zoals vergroting van de professionaliteit van de spelleidsters en het gaan werken met tweetalige leidsters, naast nog vele andere aanbevelingen (Groenendaal & Kurvers, 1992).

Het KOST-programma wordt ontwikkeld door een samenwerkingsverband van de vier grote gemeenten (de Stichting Vier Steden Project Welzijn). Het is in tegenstelling tot de meeste andere programma's een centre-based stimuleringsprogramma dat gebaseerd is op de ontwikkelingstabellen van Beller (1984). De 'centres' zijn de peuterspeelzalen en de kinderdagverblijven. Op basis van het ontwikkelingsprofiel dat per kind uit de tabellen wordt afgeleid, wordt een stimuleringsprogramma opgesteld dat aansluit op de eigen, spontane activiteit van het desbetreffende kind. Evaluatiegegevens met betrekking tot het KOST-programma zijn niet beschikbaar.

In het kader van het onderwijsvoorrangsbeleid zoals hieraan door de onderwijsvoorrangsgebieden in Nederland vorm wordt gegeven ontstaat eveneens meer aandacht voor de voorschoolse periode (in dit geval nul tot vier jaar). Verhaak en Kristensen (1992) hebben een inventarisatie-onderzoek verricht onder de zeventig onderwijsvoorrangsgebieden; de meeste voorschoolse activiteiten van deze gebieden zijn geconcentreerd op kinderen vanaf 2,5 à drie jaar. Jongere kinderen komen niet of nauwelijks met activiteiten van de onderwijsvoorrangsgebieden in aanraking. De wat oudere kinderen krijgen vooral te maken met boek- en spelpromotie (via bibliotheken, kinderdagverblijven en peuterspeelzalen, basisscholen en sociaal-cultureel werk). Via de zojuist genoemde instanties worden ook nogal wat activiteiten aangeboden op het gebied van Nederlands-als-tweede-taal. De resultaten van al deze activiteiten moeten vooralsnog als 'vaag' (Verhaak & Kristensen, 1992; p. 81) omschreven worden; zij blijken nergens getoetst te zijn. Hetzelfde kan geconcludeerd worden op basis van de inventarisatie van het onderzoek naar jeugdige allochtonen sinds 1989 (Van der Veen & Pels, 1993). Uit deze inventarisatie blijkt dat het aantal onderzoekingen dat zich daadwerkelijk met de evaluatie van programma's voor nul- tot zesjarigen bezighoudt, zeer gering is.

Veen en Vermeulen (1993) hebben een studie verricht naar binnen- en buitenlandse activiteiten gericht op de taalstimulering van meertalige kinderen in kindercentra. Behalve allerlei Nederlandse initiatieven bespreken zij ook ini-

tiatieven uit Zweden, Duitsland, de Verenigde Staten en Canada, evenals (zij het in beperkte mate) uit Engeland, Frankrijk en België. Doordat de meeste van deze (vaak experimentele) initiatieven nauwelijks van degelijke onderzoeksresultaten zijn voorzien, kunnen zij moeilijk op hun merites beoordeeld worden (p. 98).

Effectief?

De ontwikkelingen die in deze kroniek geschetst zijn geven duidelijk de nog steeds toenemende belangstelling voor de voorschoolse periode (hoe dan ook gedefinieerd) aan. Niettemin is er een aantal kritische vragen over te stellen. De zinvolheid van een dergelijke vroegtijdige stimulering wordt voornamelijk gebaseerd op allerlei ontwikkelings-psychologische overwegingen, die zijn samen te vatten in de uitspraak 'Hoe eerder stimuleren, hoe beter'. Klopt die uitspraak wel? Welke empirische gegevens zijn daarvoor aanwezig?

Een tweede vraag betreft de efficiency van dit type activiteiten, aangenomen dat zij effectief zijn tenminste; gezinsactiveringsprogramma's met al dan niet professionele huisbezoeksters zijn arbeidsintensief, en dus duur. Kunnen dezelfde effecten niet bereikt worden met minder kosten, bijvoorbeeld via centre-based activiteiten? En kan dat 'centre' dan niet een bestaande instelling zijn, zoals een peuterdagverblijf respectievelijk de groepen 1 en 2 van het basisonderwijs? Belangrijk is verder nog dat in bestaande instellingen de participatiegraad (van de kinderen) optimaal is, terwijl bij gezinsactiveringsprogramma's uitval respectievelijk het niet benaderbaar zijn van de ouders niet ongebruikelijk is (cf. Eldering & Vedder, 1992).

Allereerst: de effecten van voorschoolse programma's. In dit geval van programma's voor 0- tot 4-jarigen en, voor de buiten het reguliere onderwijs plaatsvindende activiteiten, voor vier- tot zesjarigen. De meest uitvoerige analyse van de effecten van dit type programma's is gemaakt door Farran (1990). Hij analyseerde de effecten van 32 programma's, waaraan ruim 5.000 leerlingen deelnamen. Van deze programma's waren er tien home-based en 21 centre-based. Eén programma beide. Zijn conclusie luidde dat er geen belangrijke langetermijneffecten van deze programma's vastgesteld konden worden. Hij wijst er bovendien op dat dezelfde en zelfs betere effecten bereikt kunnen worden door uitsluitend verbeteringen in het lager onderwijs aan te brengen. Over de positieve evaluatieresultaten van het High/Scope Preschool Curriculum van Weikart (zie de OpStap-paragraaf) merkt Farran op dat deze in het gebied liggen van niet door het programma bedoelde effecten. Zijn grootste problemen liggen echter op methodologisch gebied: te kleine steekproeven, niet-aselecte steekproeven, standaarddeviaties die groter zijn dan de gemiddelden (!) en zelfrapportages van de proefpersonen over hun sociaal gedrag e.d. op basis van niet-gevalideerde vragenlijsten, waarvan ook de betrouwbaarheid onbekend is (p. 511). De voornaamste conclusie van Farrans analyse is dat 'hoe vroeger stimuleren hoe beter' door de evaluatiegegevens van 'vroege' stimuleringsprojecten niet ondersteund kan worden. Het door Farran geschetste beeld past bij dat van de OpStap-evaluatie van Eldering en Vedder, evenals bij de uitkomsten van eerdere Nederlandse evaluaties van de effecten van voorschoolse programma's (cf. Slavenburg & Peters, 1989). Kortom, geen beeld om vrolijk van te worden.

Farran (1990) noemt een aantal onderzoekingen die verbeteringen in het lager onderwijs hebben aangebracht die leidden tot ten minste zulke goede resultaten als de voorschoolse programma's, zo niet tot betere. Dergelijke resultaten worden eveneens regelmatig door bijvoorbeeld Slavin, Karweit en Madden (1989) en Slavin, Madden, Karweit, Livermon en Dolon (1990) gerapporteerd, terwijl er ook in het effectieve-scholenonderzoek voorbeelden van te vinden zijn (Van der Werf & Weide, 1991), overigens met alle bezwaren die aan dit laatste type onderzoek kleven. Wij moeten daarmee betrekkelijk voorzichtig omgaan. Het is, zo moeten wij concluderen, nog lang geen uitgemaakte zaak dat voorschoolse stimulering een wezenlijke bijdrage levert aan de verbetering van de onderwijskansen van kinderen uit achterstandssituaties.

In dit verband is het verder relevant te wijzen op de mogelijkheid dat centre-based programma's (bijvoorbeeld in kinderdagverblijven) niet alleen het probleem van het uitvallen

van ouders oplossen (zij zijn niet meer de centrale beïnvloeders), maar wellicht net zo effectief kunnen zijn als home-based programma's. In een omvangrijke meta-analyse van ruim 40 onderzoeken waarin de rol van de ouders als uitvoerders van een activeringsprogramma, aanvullend op een centre-based programma, is uitgezocht komen White, Taylor en Moss (1992) tot de conclusie dat het toevoegen van een oudercomponent aan een centre-based programma geen vermeerdering van effect te zien geeft.

De ontwikkelingspsychologische (en ook vele andere) overwegingen (cf. bijvoorbeeld Meisels & Shonkoff, 1990, deel III) om vroegtijdig te beginnen met het door middel van voorschoolse programma's stimuleren van de ontwikkeling van kinderen uit sociaal-economisch lagere milieus, ondanks de niet rooskleurige resultaten van de voorschoolse programma's tot nu toe, blijven uiteraard van kracht. De politieke en maatschappelijke realiteit gebiedt bovendien dat voortgegaan wordt op de ingeslagen weg. Anderzijds zou er geprobeerd moeten worden om ten aanzien van de te realiseren langere-termijnresultaten, een stap voorwaarts te zetten. Daarvoor is echter, naar onze mening, een andere aanpak van de problematiek noodzakelijk: het oprichten van een instituut waarin continu aan het ontwikkelen èn gestaag verbeteren van programma's voor leerlingen uit sociaal-economisch lagere groepen gewerkt kan worden. Ten aanzien van ontwikkeling en evaluatie zou het toch enigszins van toeval afhankelijke beleid verlaten moeten worden. Geprobeerd zou moeten worden een aan eerder bedoeld instituut gekoppelde proefsituatie te creëren voor programma-ontwikkeling en -evaluatie waar via, hoofdzakelijk, trial-and-error geprobeerd wordt om meer effectieve programma's samen te stellen. Dat kost vele jaren. Een parallel met het ontwikkelen van andere zaken ligt voor de hand. Medicijnen worden in het algemeen ontwikkeld in zeer langdurige projecten (soms van vele tientallen jaren), datzelfde geldt voor allerlei geavanceerde technologische apparatuur. Vergeleken met het wijzigen van menselijk gedrag, in dit geval tijdens de vroegkinderlijke ontwikkeling, zijn dat bovendien betrekkelijk eenvoudige problemen. En om die op te lossen zijn speciale research- en ontwikkelingsinstituten opgericht, die over veel geld en zeer veel tijd beschikken. Wij verwachten dat een dergelijk instituut, ook met veel geld en veel tijd, eveneens nodig is om een belangrijke stap voorwaarts te zetten in de bestrijding van de onderwijsachterstandenproblematiek. Tot het zover is zullen we het onder meer met de voorschoolse programma's moeten doen, die eigenlijk nog te weinig effecten sorteren.

J. H. Slavenburg, Centrum Educatieve Dienstverlening, Rotterdam

Literatuur

Allochtone Kleuters Meer In Actie (1993). Een advies voor het doen uitvoeren van een tweetal experimenten in de voorschoolse periode ten behoeve van WVC en O&W. Amsterdam: Averroèsstichting.

Barnett, W. S. (1985). *The Perry Preschool Program and its long-term effects. A benefit-cost analysis.* Ypsilanti: High/Scope Educational Research Foundation.

Bekkers, B., Embricqs, R. van, & Loggem, D. van (1992). De gezinsactiveringsprogramma's Opstap, Opstapje en Klimrek. In J. Rispens & B. F. van der Meulen (red.), *Gezinsgerichte stimulering van kinderen in achterstandssituaties.* Amsterdam: Swets & Zeitlinger.

Beller, E. K. (1984). Die pedagogische Qualitizierung von Krippenbetreuern – eine empirische Untersuchung. In Th. Hellbrugge (Ed.), *Kinderkrippen - Krippenkinder.* Urban und Schwanenberg.

Eldering, L., & Choenni-Gobardhan, S. (1990). *Interventieprogramma's voor jonge kinderen uit kansarme groepen.* Leiden: DSWO Press.

Eldering, L., & Vedder, P. (1992). *OPSTAP. Een opstap naar meer schoolsucces?* Amsterdam: Swets & Zeitlinger.

Farran, D. C. (1990). Effects of intervention with disadvantaged and disabled children: a decade review. In S. J. Meisels & J. P. Shonkoff (Eds.), *Handbook of early childhood intervention.* Cambridge: Cambridge University Press.

Felix, Ch. (1992). *Waardering van het OPSTAPProgramma.* Amsterdam: SCO.

Groenendaal, J., & Kurvers, J. (red.) (1992). *Op het spel.* Rijswijk: Ministerie van WVC.

Jansen-Vos, F. (1990). *Basisontwikkeling.* Assen: Van Gorcum.

Leseman, P. (1989). *Structurele en pedagogische determinanten van schoolloopbanen*. Rotterdam: Rotterdamse School Advies Dienst.

Lombard, A. D. (1981). *Success begins at school*. Toronto: Lexington Books.

Madden, J., O'Hara, J., & Levenstein, O. (1984). Home again: effects of the Mother-Child Home Program on mother and child. *Child Development, 55*, 636-647.

Meisels, S. J., & Shonkoff, J. P. (1990). *Handbook of early childhood intervention*. Cambridge: Cambridge University Press.

Op-Stap-Opnieuw, Cie. (1992). *Op-Stap-Opnieuw. Voorstel voor de ontwikkeling van een nieuw gezinsgericht stimuleringsprogramma*. Rijswijk: Ministerie van WVC.

Pels, T. (red.) (1990). *Op-Stap onder de loep*. Rijswijk: Ministerie van WVC.

Riksen-Walraven, J. M. A. (1992). Instapje: de ontwikkeling van een interventieprogramma voor ouders met kinderen van 12 tot 18 maanden. In J. Rispens & B. F. van der Meulen (red.), *Gezinsgerichte stimulering van kinderen in achterstandssituaties*. Amsterdam: Swets & Zeitlinger.

Rispens, J., & Meulen, B. F. van der (1992). *Gezinsgerichte stimulering van kinderen in achterstandssituaties*. Amsterdam: Swets & Zeitlinger.

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

Slavin, R. E., Karweit, N. L., & Madden, M. A. (1989). *Effective programs for students at risk*. Needham Heights: Allyn & Bacon.

Slavin, R. E., Madden, M. A., Karweit, N. L., Livermon, B. J., & Dolon, L. (1990). Success for All: First-Year Outcomes of a comprehensive plan for reforming urban education. *American Educational Research Journal, 27*, 255-278.

Vallen, T. (red.) (1992). *Op-Stap onder vuur?* Rijswijk: Ministerie van WVC.

Veen, G. van der, & Pels, T. (1993). *Onderzoek naar jeugdige allochtonen*. Leiden: PEWA, Faculteit Sociale Wetenschappen.

Veen, A. M., & Vermeulen, A. C. A. M. (1993). *Kindercentra en meertaligheid*. Amsterdam: SCO-Kohnstamm Instituut.

Verhaak, C., Kristensen, D. (1992). *Projecten voor jonge allochtone kinderen*. Rijswijk: Ministerie van WVC.

Werf, M. P. C. van der, & Weide, M. G. (1991). Effectief onderwijs voor allochtone leerlingen. *Tijdschrift voor Onderwijsresearch, 16*, 231-243.

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.

Boekbespreking

M. van Tulder

Nascholing en onderwijsvernieuwing

Proefschrift Katholieke Universiteit Nijmegen 1992
IX + 174 pagina's ISBN 90 900 4951 7

Van Tulder heeft in haar proefschrift de nascholing voor leraren in het basisonderwijs onderzocht. Zij heeft deze nascholing in het kader van de onderwijsvernieuwing geplaatst, zij liet zich leiden door de vraag of nascholing tot vernieuwing van het basisonderwijs leidt. De overheid ziet nascholing als een belangrijk instrument voor de vernieuwing van het onderwijs. Er bestaan echter weinig empirische gegevens over de inrichting en de effectiviteit van de nascholing. Van Tulder heeft met de drie in haar proefschrift gepubliceerde onderzoeken geprobeerd daarin verandering te brengen.

Allereerst heeft zij geprobeerd vast te stellen welke kenmerken effectieve nascholing zou moeten bezitten. Zij heeft daartoe de onderzoeksliteratuur doorgenomen en daaruit een groot aantal kenmerken afgeleid die zij, via een Delphi-procedure, aan experts ter beoordeling heeft voorgelegd. Deze beoordeling leidde tot een lijst van kenmerken waaraan effectieve nascholing zou moeten voldoen. Daarna is Van Tulder, via een survey onder alle betrokkenen, nagegaan in welke mate nascholingscursussen uit de jaren 1984 t/m 1987 aan deze effectiviteitskenmerken voldeden en tot welke resultaten deze cursussen hebben geleid, waarbij de kenmerken aan deze resultaten gerelateerd werden. Ten slotte heeft zij, via gevalstudies, geprobeerd condities op schoolniveau op het spoor te komen die de resultaten van de nascholing positief beïnvloed hebben.

Delphi-procedure

Een panel van 36 experts beoordeelde in drie rondes een lijst van kenmerken van effectieve nascholing zoals die uit de onderzoeksliteratuur waren afgeleid. Uiteindelijk ontstond overeenstemming over 34 kenmerken. Het gaat te ver om deze kenmerken hier op te sommen. Ze liggen op het terrein van programmering en voorbereiding, opzet, uitvoering, implementatiekansvergroting, randvoorwaarden en kenmerken van nascholers.

Survey

Door middel van interviews bij en enquêtes onder de betrokkenen bij de nascholingscursussen in het kader van het innovatieproces basisonderwijs in de jaren 1984-1987, heeft Van Tulder onderzocht in welke mate nascholingscursussen voldeden aan de 34 kenmerken van effectieve cursussen. Grosso modo bleek (dit is interpretatie recensent) dat in ongeveer éénderde van de gevallen aan een aantal van deze kenmerken voldaan was.

De effecten van de nascholing waren eveneens onderwerp van studie in het survey. De leraren zowel als de schoolleiding waren van mening dat deze gering waren; de kans op blijvend gebruik van het geleerde werd matig geacht. Wel was men in het algemeen tevreden over de cursus.

De relatie tussen de kenmerken van de nascholingscursussen en de door de betrokkenen gepercipieerde effecten werd vastgesteld d.m.v. correlatie- en regressie-analyse. Er werden nogal wat kenmerken gevonden die een positief verband hadden met deze effecten. Voor het effect 'daadwerkelijk gebruik van de vernieuwing' waren vooral van belang: praktijkgerichtheid, en buiten de cursussen door de leerkrachten extra geïnvesteerde tijd. Ook enkele andere kenmerken bleken de kans op implementatie van de vernieuwing te vergroten: o.a. onderlinge informatie-uitwisseling tussen leerkrachten en teamgerichte in plaats van individuele nascholing.

Gevalstudies

Voor de gevalstudies zijn 6 schoolteams geselecteerd die enerzijds een nascholingscursus gevolgd hadden die in redelijke tot in hoge mate voldeed aan de kenmerken van effectieve nascholing en die anderzijds verschilden in effect van de nascholing: 4 scholen gekenmerkt door 'groot effect' en 2 door 'weinig'. Op scholen met grote effecten bleken de nascholingscursussen vooral gevolgd te zijn vanwege specifieke problemen en behoeften, waren voorzieningen getroffen om de implementatie te vergemakkelijken en was de cursus aangepast aan de behoeften en ervaren problemen bij de implementatie van de innovatie. Op deze scholen vonden planning en coördinatie van het implementatieproces plaats, werden ervaringen uitgewisseld en werd ondersteuning gegeven; evaluatie- en voort-

gangscontrole werden eveneens gerealiseerd.

Van Tulder komt op basis van haar onderzoeksresultaten tot een aantal aanbevelingen voor de vormgeving van nascholingscursussen, waarbij zij rekening probeert te houden met de nieuwe besturingsfilosofie volgens welke scholen nascholing 'kopen' bij de nascholer die hun goeddunkt. Dat kan volgens Van Tulder alleen goed gebeuren als de scholen over een beleids- of ontwikkelingsplan beschikken waarin zij de na te streven vernieuwingen hebben vastgelegd. Op basis van dat plan kunnen zij beoordelen welke nascholing relevant is. Tot op heden, zo stelt zij, hebben scholen hun nascholingsbehoeften niet helder vastgelegd. Zij hoopt dat haar onderzoek richtlijnen geeft aan de scholen voor de wijze waarop nascholing in het kader van de onderwijsvernieuwing benut kan worden.

Het is vrijwel ondoenlijk de inhoud van de onderzoeken van Van Tulder beknopt samen te vatten. De zojuist gegeven samenvatting is maar een zwakke afspiegeling van al hetgeen zij aan de orde heeft gesteld. En zij niet alleen; het proefschrift bevat artikelen die zij samen met andere auteurs (onder wie S. Veenman, M. Voeten en R. van der Vegt) heeft geschreven. In de Nederlandse maar ook in de buitenlandse literatuur is het onderwerp 'nascholing' èn de effectiviteit ervan niet erg populair. Er wordt relatief weinig aandacht aan besteed en als dat al gebeurt, bestaat de effectiviteitsbepaling veelal uit door de cursisten gerapporteerde effecten; minder subjectieve metingen worden veelal niet gehanteerd. Gelet op de geringe aandacht voor dit onderwerp is het proefschrift van Van Tulder een grote aanwinst; gelet op de wijze waarop de effecten gemeten zijn past het (helaas) bij de onderzoekstraditie op dit terrein. Uit haar gevalstudies heeft Van Tulder de indruk overgehouden dat er feitelijk meer geïmplementeerd was dan door de betrokkenen was gerapporteerd. Een argument van een meer objectieve effect-bepaling dus!

De auteur beschrijft haar onderzoeken zeer nauwgezet en zeer gedetailleerd. Dat bemoeilijkt het trekken van de meer globale conclusies waaraan praktici nu eenmaal behoefte hebben. Wetenschappelijke nuancering is natuurlijk uitstekend, praktische bruikbaarheid is toch wat anders.

Terecht wijst Van Tulder op het doorslaggevende belang van de transfer naar de onderwijssituatie van hetgeen in de nascholing is geleerd; terecht stipt zij in dit verband ook de rol van 'coaching' aan (in Nederland meestal aangeduid met de term 'klasse-consultatie'). In haar gevalstudies somt zij allerlei activiteiten op die van belang zijn voor de feitelijke invoering van een onderwijsvernieuwing, maar die in Nederland (met zijn kleine basisscholen) meestal door schoolbegeleiders worden uitgevoerd (evenals de zojuist genoemde 'coaching'). Het zijn m.n dit soort schoolbegeleidende activiteiten die de transfer kennelijk bevorderen. Deze opmerking brengt mij bij mijn voornaamste punt van kritiek op de dissertatie van Van Tulder: zij beschouwt nascholing kennelijk als een belangrijk onderwijsinnovatie-instrument, zonder een vergelijking te maken met andere daarvoor geschikte instrumenten. Het onderzoek tot nu toe, ook dat van haarzelf, toont echter het omgekeerde aan: nascholing is geen belangrijk instrument voor onderwijsvernieuwing omdat het slechts geringe effecten sorteert. En die effecten zullen, naar mijn verwachting, nog verder afnemen nu de scholen zelf gaan inkopen omdat daarbij allerlei niet voor de innovatie relevante factoren zullen gaan meespelen.

Al met al heeft Van Tulder een belangrijk, zeer goed onderbouwd (maar soms wel erg gedetailleerd en genuanceerd) proefschrift geschreven over een nog goeddeels braakliggend terrein, waarin zij naar mijn mening nascholing echter als een te belangrijk onderwijsinnovatie-instrument beschouwt.

J. H. Slavenburg

Mededelingen

Symposium Gezinsonderzoek

Het Institute of Family Studies van de Katholieke Universiteit Nijmegen organiseert op 25 en 26 november 1993 een symposium met als thema 'Pedagogische advisering en begeleiding', in het kader van Kinderbescherming, Jeugdhulpverlening, Preventie en Gezinsondersteuning.

Voor nadere inlichtingen en inschrijving kunt u contact opnemen met de organisatie van het symposium: R. van Zutphen, mevr. D. van Boxtel of mevrouw M. Bekker, Institute of Family Studies, Postbus 9103, 6500 HD Nijmegen, tel.: 080 - 612144/612955/612293.

International Standing Conference for the History of Education History of Instruction, Education and Cultural Transfer since the XVIth century. Amsterdam, August 10-13, 1994

Please let us know if you are interested in coming and/or presenting a paper. Address: Conference service Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam.

SIG Writing Conference

The Special Interest Group WRITING of the European Association for Research on Learning and Instruction (EARLI) invites researchers to participate in and contribute to the SIG Writing Conference, October 20th-21th, 1994.

The Conference will be organized by the Graduate School of Teaching and Learning of the University of Amsterdam and the Centre for Language and Communication of the Utrecht University. The organizers aim to organize a conference with formal lectures. Peer review selections of the contributions will be published. On behalf of the SIG Writing we invite you to send us your contributions. Proposed themes are: 1) What is writing? 2) Writing in education; 3) Effective instruction; 4) Developmental research in writing; 5) Methodology in writing research.

Further information: H. van den Bergh, Centre for Language and Communication, University of Utrecht, Trans 10, 3512 JK Utrecht, tel.: 030 - 538137 or G. Rijlaarsdam, Graduate School of Teaching and Learning, University of Amsterdam, Herengracht 256, 1016 BV Amsterdam, tel.: 020 - 5253541.

Inhoud andere tijdschriften

Tijdschrift voor Onderwijsresearch
18e jaargang, nr. 1, 1993

Mijn kruistabel heeft zo veel verwachte waarden kleiner dan 5. Wat moet ik nu doen? door P. M. Kroonenburg en A. Verbeek
Een onderzoek naar betrouwbaarheid en validiteit van studentoordelen over de didactische vaardigheid van universitaire docenten, door E. van Schooten, M. Eiting en T. M. Bechger
Multiple correspondentie-analyse van school- en beroepsloopbanen, door T. W. Taris en P. G. M. van der Heijden
De veronderstelde bekwaamheid als voorspeller van studieresultaten, door A. Vrugt, M. Ph. Langereis en J. Hoogstraten
Notities en Commentaren
'Transfer: onderzoek en onderwijs'. Een reactie op de onderliggende denkopvatting, door B. Wolters
Hoe soms is vaak? door Ch. G. J. Michels en M. J. M. Voeten

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 6, 1993

Basisdiagnostiek in de psychosociale jeugdhulpverlening: de Vragenlijst Sociale en Pedagogische Situatie, door E. M. Scholte
Onderzoek naar de perspectiefnemingsvaardigheden van kinderen met autisme, door H. Roeyers en K. Impens
Doen of laten? door C. Verwers en P. H. van der Laan

Rectificatie

Betreft: Pedagogische Studiën, 70/3, pagina 195

Tot onze spijt is de tekst van de Samenvatting behorende bij het artikel 'De invloed van het gebruikte leerboek op het maken van specifieke

fouten bij wiskundeproblemen' van J. Chr. Perrenet enigszins verminkt afgedrukt.
Onderstaand treft u de volledige tekst aan.

Samenvatting

Wiskundemethoden verschillen in didactische opzet. In onderzoek van De Leeuw, Meijer, Perrenet en Groen (1988) met een zogeheten transfertest werd tussen drie methoden geen verschil gevonden op prestaties van vwo-leerlingen aan het einde van het derde leerjaar. Dit artikel beschrijft een nadere analyse van dat onderzoeksmateriaal op het voorkomen van enkele specifieke fouten in de schriftelijke uitwerkingen door de leerlingen van de transfertestopgaven. Uit methodekenmerken worden hypothesen afgeleid betreffende twee typen fouten: rekenfouten en wiskundetaal-fouten (fouten in notatie en terminologie). Bij beide fouttypen wordt na controle voor de invloed van de factor wiskundige vaardigheid en de factor leraar geen verschil gevonden. De leerlingen van de methode met vanaf het begin de meeste nadruk op formele wiskundetaal maken tegen de verwachting in meer wiskundetaal-fouten, maar het verschil is niet significant. De rol van formele wiskundetaal in het wiskunde-onderwijs wordt besproken in relatie tot het nieuwe wiskundeleerplan voor de onderbouw.

TS oct 6441

Inhoudsopgave

Pedagogische Studiën Jaargang 70, nummer 5/1993

R. Damhuis en K. de Glopper
Leerkrachtoordelen over de wenselijke woordenschat van 4- tot 6-jarigen: betrouwbaarheid, validiteit en gebruiksmogelijkheden 322

E. G. Harskamp
Spelen op school 337

E. Roelofs, S. Veenman en J. Raemaekers
Teamgerichte nascholing en coaching: Effecten op het didactisch handelen van leerkrachten in combinatieklassen 352

H. M. Y. Koomen, J. B. Hoeksma en M. Meerum Terwogt
Het belang van informatie over een afwezige genetische vader 368

Discussie
A. M. P. Knoers
De autonomievergroting van scholen. Zin en onzin 382

Kroniek
P. J. J. Stijnen (red.)
De Onderwijsresearchdagen 1993 in Maastricht: 26-28 mei 387

Mededelingen 399

BIBLIOTHEEK DER
RIJKSUNIVERSITEIT
UTRECHT

Leerkrachtoordelen over de wenselijke woordenschat van 4- tot 6-jarigen: betrouwbaarheid, validiteit en gebruiksmogelijkheden*

R. Damhuis en K. de Glopper**

Samenvatting

In het kader van de ontwikkeling van *Taalplan Kleuters*, een taalstimuleringsmodel voor het onderwijs Nederlands aan kleuters in meertalige groepen, zijn middels een schriftelijke vragenlijst leerkrachtoordelen verzameld over het moment waarop kleuters bepaalde woorden dienen te beheersen. De 2585 beoordeelde woorden vormen de basis van Taalplan Kleuters. Op basis van de leerkrachtoordelen is een ordening van deze woorden binnen de twee kleuterjaren tot stand gebracht. In dit artikel worden de betrouwbaarheid en validiteit van de oordelen van de kleuterleerkrachten en de gebruiksmogelijkheden van de geordende woordenlijst besproken. De absolute oordelen van de leerkrachten verschillen naar gelang het percentage anderstalige kleuters op school. Wat betreft de relatieve oordelen komen de groepen leerkrachten sterk overeen. De leerkrachtoordelen hangen nauw samen met de feitelijke woordkennis van leerlingen. Als streefdoelstelling lijkt de woordenlijst niet te laag te mikken.

1 Inleiding

Op het ogenblik is bijna 7% van de leerlingen in het basisonderwijs van niet-Nederlandse herkomst (Vallen & Stijnen, 1991). Veel van deze anderstalige kinderen vertonen vergeleken met van huis uit Nederlandstalige leeftijdsgenoten omvangrijke achterstanden in hun Nederlandse taalvaardigheid. Dit geldt niet alleen voor neveninstromers, maar ook voor kinderen die vanaf hun vierde jaar al deelnemen aan het Nederlandse onderwijs. De achterstanden zijn niet voor alle etnische groepen even groot. Peilingsonderzoek onder 6-jarigen laat zien dat de achterstanden van Surinaamse en Antilliaanse leerlingen relatief klein zijn. Marokkaanse en vooral Turkse kinderen zijn beduidend verder achter. De woordenschat is een van de gebieden waarop de achterstand het grootst is (Boogaard, Damhuis, De Glopper & Van den Bergh, 1990). Grote achterstanden in de woordenschat zijn ook geconstateerd door Verhoeven en Vermeer (1985): zij constateerden voor 8-jarige meertalige leerlingen een achterstand van vier jaar op Nederlandstalige leerlingen. Hun recente overzicht van cross-sectionele toetsgegevens maakt aannemelijk dat in de loop van het basisonderwijs de verschillen in woordenschat tussen Nederlandstalige en anderstalige leerlingen relatief gezien weliswaar gelijk blijven, maar in absolute zin groter worden (Verhoeven & Vermeer, 1992).

Dat er ten aanzien van de woordenschat grote verschillen optreden is niet verbazend. Vergeleken met andere aspecten van taalvaardigheid, zoals artjculatie of zinsvorming, vormt de woordenschat een relatief ongestructureerd domein. Bovendien vormt het leren begrijpen en gebruiken van woorden een enorme taak; dit geldt voor moedertaal- of T1-sprekers, maar zeker ook voor tweede-taalverwervers. Gedurende hun schoolloopbaan komen leerlin-

* Het hier gerapporteerde onderzoek maakte deel uit van het project *Intensivering Taalplan Kleuters*, dat gesubsidieerd is via de Projectgroep NT2 van het Ministerie van Onderwijs en Wetenschappen. Het project sloot aan bij het reeds lopende project *Taalplan Kleuters* van Het Projectbureau (Rotterdam).
** De auteurs bedanken M. Boers en M. Kienstra van Het Projectbureau (Rotterdam), alsmede M. Eiting van het SCO-Kohnstamm Instituut voor hun bijdrage aan het totstandkomen van dit artikel.

gen jaarlijks met vele duizenden onbekende woorden in aanraking. Incidenteel leren speelt bij de woordenschatverwerving een belangrijke rol. Volgens schattingen resulteert tussen 25 en 50% van de aanwas van het receptief vocabulaire uit de confrontatie met geschreven tekst (Nagy & Herman, 1987). Zeer veel woorden worden buiten, maar ook binnen de school zonder expliciete instructie uit de context geleerd (Anderson, Wilson & Fielding, 1988; Nagy & Anderson, 1984; Sternberg, 1987). Vocabulaire wordt natuurlijk ook binnen het woordenschatonderwijs geleerd, via directe instructie van de betekenis van specifieke woorden, maar niet alle woorden die in het taalaanbod voorkomen, kunnen door intentioneel leren verworven worden (Jenkins, Matlock & Slocum, 1989). Bij het leren van de betekenis van onbekende woorden uit contexten zijn verschillende vaardigheden in het geding, waaronder kennis van de morfologische en syntactische eigenschappen van woorden en vaardigheden in het afleiden van woordbetekenissen uit de talige context. Ook de achtergrondkennis van taalgebruikers is daarbij van belang (Drum & Konopak, 1987; Sternberg, 1987).

Leerlingen verschillen sterk in de mate waarin zij beschikken over de vaardigheden die bij het leren van woorden in het geding zijn (De Glopper, Van Daalen-Kapteijns & Schouten-Van Parreren, 1992; Schouten-Van Parreren, De Glopper & Van Daalen-Kapteijns, 1993). Verder geldt dat uitbreiding van het vocabulaire afhankelijk is van de reeds aanwezige woordkennis. De woorden die een taalverwerver kent, fungeren als kapstok voor het afleiden, aanleren en onthouden van betekenissen van nieuwe woorden. En naarmate meer woorden bekend zijn, is het aantal aanknopingspunten voor nieuwe woorden groter.

Gezien het belang van woordkennis en de aanzienlijke achterstanden van anderstalige leerlingen dient in het onderwijs aan anderstaligen veel aandacht aan de ontwikkeling van de woordenschat besteed te worden. De vraag hoe de woordenschatontwikkeling precies gestimuleerd moet worden, is echter niet direct te beantwoorden. Aan deze vraag is namelijk een reeks deelvragen verbonden, waaronder de volgende: Welke woorden moeten aangeboden worden? Op welk moment? Aan welke leerlingen? Op welke wijze moeten woorden geïntroduceerd worden? Hoe moet het onthouden van nieuwe woorden bevorderd worden? Hoe kunnen leerlingen zelfstandig nieuwe woorden leren?

Uit onderzoek naar de inrichting van woordenschatonderwijs blijkt dat leerkrachten voornamelijk improviserend werken, zonder systematische aanpak. Als voornaamste probleem wordt gemeld het ontbreken van een concrete leerlijn voor de hele basisschool waarin woordenschat systematisch wordt opgebouwd en uitgebreid (Verhallen, 1991).

Het Taalplan Kleuters (Boers, Kienstra & Padmos, 1993; Boers & Kienstra, 1993) is een taalstimuleringsmodel voor gebruik in meertalige kleutergroepen waarin getracht wordt de stimulering van de woordenschatontwikkeling van anderstalige leerlingen zo vroeg mogelijk te starten. De kernleerlijn in Taalplan Kleuters betreft de woordenschat. Deze leerlijn is gebaseerd op de Unaniemenlijst (Kohnstamm, Schaerlaekens, De Vries, Akkerhuis & Froonincksx, 1981) en de produktielijsten van Coenen en Vermeer (1988). Omdat deze bronnen geen informatie boden over de wenselijke ordening van woorden binnen de kleuterperiode, is een vragenlijstonderzoek onder leerkrachten uitgevoerd. Op basis van dit vragenlijstonderzoek is een rangordelijst voor de wenselijke receptieve beheersing van de woorden samengesteld (Damhuis, De Glopper, Boers & Kienstra, 1992): een woordenschatstreeflijst die geldt voor jonge kinderen, ongeacht hun herkomst. Het onderhavige artikel doet verslag van het uitgevoerde vragenlijstonderzoek. Twee vragen staan hierbij centraal: Hoe betrouwbaar en valide zijn de oordelen die de leerkrachten in de vragenlijst hebben gegeven? Op welke wijze kunnen de leerkrachtoordelen gebruikt worden voor de inrichting van het woordenschatonderwijs aan kleuters?

2 Methode van onderzoek

2.1 Samenstelling van de woordenlijst

In het project Taalplan Kleuters is in de loop van enkele schooljaren lesmateriaal Nederlands als tweede taal voor meertalige kleutergroepen ontwikkeld, in nauwe samenwerking met scholen en leerkrachten. Voor het schooljaar 1989 - 1990 is een *Voorlopige woordenlijst*

Taalplan Kleuters samengesteld. De basis voor deze woordenlijst vormden twee bestaande woordenlijsten voor jonge kinderen, de *Nieuwe streeflijst woordenschat* (Kohnstamm e.a. 1981) en de *Nederlandse woordenschat allochtone kinderen* (Coenen & Vermeer, 1988).

De *Nieuwe streeflijst woordenschat* is gebaseerd op oordelen van leerkrachten uit de kleuter- en onderbouw van basisscholen in Nederland en België over de wenselijkheid van de receptieve beheersing van 6788 woorden. Deze woorden waren uit diverse schriftelijke bronnen afkomstig, aangevuld vanuit ervaring van en gesprekken met personen die met jonge kinderen te maken hadden (Kohnstamm e.a. 1981, pp. 5-7). Van deze lijst bleken er 1838 woorden te zijn waarvan 90% of meer van de Nederlandse leerkrachten aangaf dat ze door 6-jarigen begrepen moeten worden: de Unaniemenlijst Nederland. Deze Unaniemenlijst is opgenomen in de *Voorlopige woordenlijst Taalplan Kleuters*.

De *Nederlandse woordenschat allochtone kinderen* is gebaseerd op audio-opnamen van gesproken taal van allochtone kinderen in de leeftijd van vijf tot dertien jaar. Op basis van de door deze kinderen gebruikte woorden is een beheersingslijst samengesteld. Daarnaast zijn thematische lijsten opgenomen van woorden die meer dan één keer voorkomen. Deze thematische lijsten zijn uitgesplitst naar kleuterbouw (groep 1 en 2), onderbouw (groep 3 en 4), middenbouw (groep 5 en 6) en bovenbouw (groep 7 en 8). Uitgangspunt bij ons gebruik van deze produktieve lijsten was dat woorden die in de kleuter- en onderbouw van de basisschool door kinderen produktief beheerst worden, voorafgaand aan die periode, dus in de kleuterbouw, aangeboden moeten worden.

In de produktielijsten van Coenen en Vermeer hebben slechts 644 woorden uit het corpus voor de gehele basisschool een frequentie van 40 of meer (de betrouwbaarheidsgrens). Selectie vanaf de betrouwbaarheidsgrens voor alleen de deelcorpora kleuter- en onderbouw zou een voor ons doel veel te klein aantal woorden hebben opgeleverd. Selectie op frequentie onder de betrouwbaarheidsgrens zou resulteren in willekeur. Om deze redenen zijn alle woorden uit de kleuter- en onderbouwcorpora die in de thematische lijsten voorkomen (dus zonder hapaxen), opgenomen in de *Voorlopige woordenlijst Taalplan Kleuters*.

Ten slotte is aan de *Voorlopige woordenlijst Taalplan Kleuters* een beperkt aantal extra woorden toegevoegd, die niet in bovengenoemde twee lijsten voorkomen, maar voor kleuters binnen de thema's van Taalplan Kleuters relevant geacht werden. Een deel van deze woorden komt wel voor in de alfabetische lijst van de *Nederlandse woordenschat allochtone kinderen* bij groep 5-8.

In totaal bevat de *Voorlopige woordenlijst Taalplan Kleuters* 2585 woorden. Deze woordenlijst is binnen het kader van de ontwikkeling van het Taalplan Kleuters opgevat als een minimale streeflijst voor de woordenschatontwikkeling van kleuters. De lijst bevat de woorden die *tenminste* door kleuters beheerst zouden moeten worden, maar bevat *niet alle* woorden die kleuters beheersen of zouden moeten beheersen.

Tijdens het eerste proefinvoeringsjaar van Taalplan Kleuters (schooljaar 1989-1990) bleek dat er behoefte was aan een normatieve lijst met gegevens over zowel de receptieve als de produktieve beheersing, gedifferentieerd naar een aantal momenten binnen de kleuterperiode. De lijst van Kohnstamm e.a. (1981) is weliswaar een normatieve lijst, maar geeft alleen aan welke woorden op het *einde* van de kleuterperiode (nl. door 6-jarigen) *receptief* beheerst zouden moeten worden. De woordenlijsten van Coenen en Vermeer (1988) zijn *descriptieve* lijsten van de feitelijke *produktieve* beheersing; hun groep proefpersonen in de onderbouw van de basisschool is bovendien *vrij klein*. Ten behoeve van de verdere ontwikkeling van het Taalplan Kleuters is daarom een vragenlijstonderzoek uitgevoerd onder leerkrachten waarin de wenselijke receptieve en produktieve beheersing van de woorden uit de *Voorlopige woordenlijst Taalplan Kleuters* bevraagd zijn, uitgesplitst naar een drietal momenten binnen de kleuterperiode.

2.2 Leerkrachtensteekproef

Uit het adressenbestand van het Ministerie van Onderwijs en Wetenschappen is een landelijke, gestratificeerde steekproef van basisscholen getrokken. Uit vier strata zijn steekproeven van ieder 100 scholen getrokken:

stratum 0 = geen anderstalige kleuters;
stratum 0-10 = tussen 0 en 10 % anderstalige kleuters;
stratum 10-70 = van 10 tot 70 % anderstalige kleuters;
stratum 70+ = 70 % of meer anderstalige kleuters.

De 400 scholen zijn aangeschreven met het verzoek of een kleuterleerkracht tegen een vergoeding van een boekebon (*f* 50,-) bereid was om mee te werken aan het onderzoek. Uit de groep van 113 bereidwillige kleuterleerkrachten zijn 80 leerkrachten geselecteerd, zodanig dat de vier strata gelijk vertegenwoordigd waren en er binnen elke stratumgroep spreiding was op de volgende kenmerken: schoolgewicht, leeftijdssamenstelling van de eigen kleutergroep en het aantal jaren ervaring dat de leerkrachten in kleutergroepen hadden. De leerkrachten waren niet alleen verbonden aan scholen uit de vier grote steden, maar aan scholen uit heel Nederland. De 80 leerkrachten kregen begin mei 1990 de vragenlijst toegestuurd. Na schriftelijk en telefonisch rappel kregen we 71 ingevulde vragenlijsten terug.

2.3 Vragenlijst

De vragenlijst bestond uit een alfabetisch geordende versie van de *Voorlopige woordenlijst Taalplan Kleuters* (zie par. 2.1). De woorden zijn per pagina in twee kolommen van elk 35 woorden aangeboden. De lijst besloeg op deze manier 37 pagina's. Achter elk woord zijn antwoordkolommen gegeven met in cijfercodes de mogelijke momenten voor het begrip (receptieve beheersing) en voor het gebruik van het woord (produktieve beheersing). De leerkrachten werd gevraagd het volgens hen *wenselijke* moment voor begrip en gebruik aan te kruisen.

Voorafgaand aan de woordenlijst bevatte de vragenlijst een uitgebreide instructie voor de leerkrachten, deels afgeleid uit Kohnstamm e.a. (1981). Hoofdelementen daaruit zijn de omschrijving van het gevraagde oordeel en de geboden antwoordmogelijkheden.

De leerkrachten is gevraagd een oordeel te geven over *de woordenschat die voor een kleuter, ongeacht zijn herkomst, noodzakelijk is om aan het reguliere Nederlandse onderwijs te kunnen deelnemen.* De rangordelijst is dus een *algemene, voor alle leerlingen geldende streeflijst.* De oordelen geven het streefmoment waarop beheersing van het woord noodzakelijk wordt geacht voor deelname aan het gewone Nederlandse onderwijs. Het gaat dus om het moment waarop de beheersing *wenselijk* wordt geacht, en niet om een inschatting van de feitelijke beheersing. De leerkrachten konden uit de volgende momenten kiezen:

0: kort na binnenkomst in de kleutergroep (kleuters die pas op school zijn, moeten dit woord al vrij gauw kunnen begrijpen/zelf kunnen gebruiken);
1: aan het einde van groep 1 (kleuters die in principe het volgend schooljaar als laatste jaar in de kleutergroep zullen doorbrengen, moeten dit woord kunnen begrijpen/zelf kunnen gebruiken);
2: aan het einde van groep 2 (kleuters die in principe volgend schooljaar naar groep 3 zullen gaan, moeten dit woord kunnen begrijpen/zelf kunnen gebruiken);
3: later dan groep 2 (een kind hoeft dit woord nog niet aan het eind van de kleuterperiode zelf te kunnen gebruiken, maar pas later in de basisschool).

Voor de receptieve beheersing is code 3 niet als antwoordmogelijkheid geboden, omdat bij de ontwikkeling van het Taalplan Kleuters het uitgangspunt was dat alle woorden uit de lijst in elk geval receptief beheerst zouden moeten worden aan het einde van de kleuterperiode.

De leerkrachten werd sterk aangeraden de lijst niet in één keer te beoordelen, maar het werk te spreiden over een aantal dagen. Om toch mogelijke effecten van vermoeidheid tegen te gaan, kregen de leerkrachten een wisselende startpagina (zie Figuur 1).

	Begrip	*Gebruik*		*Begrip*	*Gebruik*
aaien	0 1 2	0 1 2 3	achterom	0 1 2	0 1 2 3
aan	0 1 2	0 1 2 3	achterop	0 1 2	0 1 2 3
aanbellen	0 1 2	0 1 2 3	achterpoot	0 1 2	0 1 2 3

Figuur 1. Fragment van pagina 1 van de vragenlijst

3 Betrouwbaarheid en validiteit van de leerkrachtoordelen

3.1 Ontbrekende waarnemingen

De leerkrachten dienden ieder 2585 woorden te beoordelen. Gezien dit grote aantal is de vraag naar het aantal ontbrekende oordelen van belang. Uit Tabel 1 blijkt dat de beoordelaars maar weinig woorden hebben overgeslagen: bij de receptieve beoordeling gemiddeld 35.3 (1.4%) en bij de produktieve beoordeling 17.8 (0.7%). Beoordelaars bij wie aanzienlijke aantallen oordelen ontbreken, hebben één of meer pagina's overgeslagen. Het hogere aantal ontbrekende waarnemingen bij de receptieve beoordeling heeft te maken met het ontbreken van de antwoordmogelijkheid 3 (zie hiervoor). Anders dan bij de produktieve oordelen konden de beoordelaars niet aangeven dat een woord nog niet aan het einde van de kleuterperiode gekend hoefde te worden. Enkele beoordelaars gaven te kennen dat zij daaraan toch behoefte hadden; zij hebben in voorkomende gevallen de betreffende woorden niet beoordeeld[1]. Over het geheel genomen lijken de beoordelaars hun taak serieus te hebben opgevat.

3.2 Effecten beginpagina

Ter bestrijding van volgorde-effecten zijn de leerkrachten op verschillende pagina's gestart met beoordelen. Per startpagina uit de vragenlijst (1, 7, 13, 19, 25 en 31) is voor de desbetreffende groep beoordelaars het gemiddelde receptieve en produktieve oordeel berekend, voor zes, met de startpagina's corresponderende delen van de vragenlijst (woordnrs. 1-420, 421-840, 841-1260, 1261-1680, 1681-2100, 2101-2585). Via de techniek van 'mean polishing' (Hoaglin, Mosteller & Tukey, 1983) zijn de gemiddelden ontleed in componenten: het algemeen gemiddelde, een jury- of beoordelaarseffect, een beginpagina-effect en een residu. Tabel 2 en 3 bevatten de gegevens voor de receptieve en de produktieve oordelen, na analyse. Per jury staat het bij de betreffende beginpagina horende residu vetgedrukt. De afzonderlijke en de gemiddelde waarden van de residuen na herordening naar beginpagina zijn weergegeven in Tabel 4. Uit deze tabel blijkt dat er zeer geringe volgorde effecten optreden. Het eerste blok woorden wordt gemiddeld genomen het minst 'streng' beoordeeld, het vierde blok het 'strengst'.

3.3 Individuele overeenstemming binnen groepen

Per stratum is de overeenstemming tussen de oordelen van de individuele leerkrachten bepaald. De gemiddelde individuele overeenstemming binnen groepen, uitgedrukt in de pmc lag voor de oordelen over de receptieve beheersing tussen 0.34 en 0.38 en voor de oordelen over de produktieve beheersing tussen 0.44 en 0.49. De overeenstemming is dus matig positief. De betrouwbaarheid, uitgedrukt in coëfficiënt α, beloopt voor de receptieve oordelen 0.90 tot 0.91, voor de produktieve beheersing 0.93 tot 0.95. De jury's zijn homogeen. De passing van een één-factormodel is steeds goed. De onder dit model geschatte Fabin-betrouwbaarheid (Hagglund, 1982) is steeds vrijwel gelijk aan α. Tabel 5 vat deze resultaten samen.

3.4 Overeenstemming tussen groepen

De overeenstemming tussen de groepen beoordelaars is bepaald via een LISREL-analyse. Per stratum zijn de beoordelaars in twee parallelle subgroepen gesplitst. Per stratum is vervolgens een ware stratum-score geschat, na gelijkstelling van de regressies op de ware score van de stratum-helften. De onder dit model geschatte correlaties tussen de ware stratum-scores belopen voor de receptieve oordelen waarden tussen .97 en .99, voor de produktieve oordelen waarden tussen .96 en .99.[2] De getoetste mo-

Tabel 1

Ontbrekende waarnemingen bij beoordeling Woordenlijst, in aantallen woorden: gemiddelde (X), standaarddeviatie (sd), 1e percentiel (P1), 1e kwartiel (Q1), mediaan (Md), 3e kwartiel (Q3) en 9e percentiel (P9)

	X	sd	P1	Q1	Md	Q3	P9
Receptieve oordelen	35.3	72.0	0	1	3	19	164
Produktieve oordelen	17.8	46.8	0	2	5	9	29

Tabel 2
Gemiddelde receptieve oordelen na ontleding: algemeen gemiddelde (AG), jury-effecten, woordgroep-effecten en residuen

Woord-nummers	Jury per beginpagina						Woordgroep-effect
	P1	P7	P13	P19	P25	P31	
1 - 420	-.03	-.02	-.01	-.01	.01	.04	.00
421 - 840	-.01	-.01	.08	.00	-.01	.03	.05
841 - 1260	-.02	-.01	.05	-.01	-.02	-.03	-.04
1261 - 2680	.00	.04	-.08	.04	.00	-.04	.02
1681 - 2100	.02	.00	-.03	-.01	.07	-.07	-.05
2101 - 2585	.02	-.03	-.03	-.02	.01	.04	.03
Jury-effect	**-.05**	**.03**	**.00**	**-.06**	**-.01**	**.11**	.80 (*AG*)

Tabel 3
Gemiddelde produktieve oordelen na ontleding: algemeen gemiddelde (AG), jury-effecten, woordgroep-effecten en residuen

Woord-nummers	Jury per beginpagina						Woordgroep-effect
	P1	P7	P13	P19	P25	P31	
1 - 420	.04	-.04	-.01	.01	.01	.05	-.02
421 - 840	-.04	.11	-.01	.01	-.10	.03	.03
841 - 1260	-.06	.08	.11	-.05	-.03	-.07	-.05
1261 - 1680	.04	-.03	-.04	.09	.00	-.06	.05
1681 - 2100	.07	-.07	-.06	-.03	.08	-.02	-.04
2101 - 2585	-.02	-.06	-.02	-.02	.05	.05	.03
Jury-effect	**.03**	**-.02**	**-.02**	**-.12**	**.04**	**.11**	1.42 (*AG*)

Tabel 4
Afzonderlijke en gemiddelde residuen na herordening naar beginpagina

	Residuen per woordgroep in volgorde beoordeling											
	Receptief						Produktief					
Jury	1e	2e	3e	4e	5e	6e	1e	2e	3e	4e	5e	6e
P1	-.03	-.01	-.02	.00	.02	.02	.04	-.04	-.06	.04	.07	-.02
P7	-.01	-.01	.04	.00	-.03	-.02	.11	.08	-.03	-.07	-.06	-.04
P13	.05	-.08	-.03	-.03	-.01	.08	.11	-.04	-.06	-.02	-.01	-.01
P19	.04	-.01	-.02	-.01	.00	-.01	.09	-.03	-.02	.01	.01	-.05
P25	.07	.01	.01	-.10	-.02	.00	.08	.05	.01	-.10	-.03	.00
P31	.04	.04	.03	-.03	-.04	-.07	.05	.05	.03	-.07	-.06	-.02
Gemidd.	.03	-.01	.00	-.03	-.01	.00	.08	.01	-.02	-.04	-.01	-.02

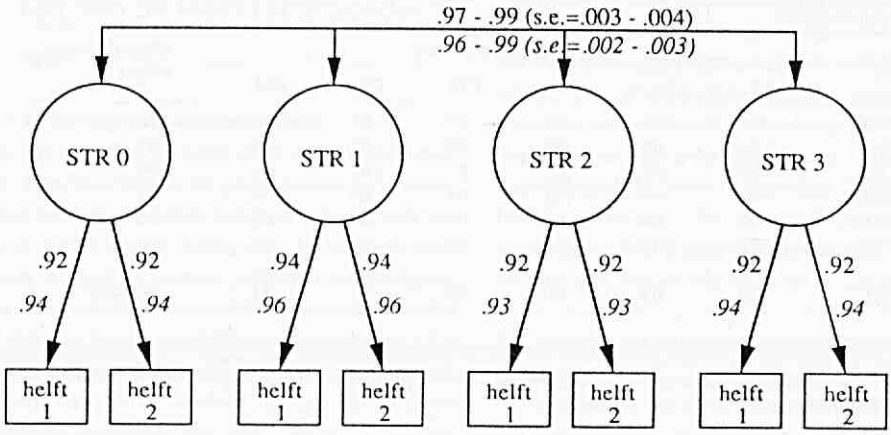
Tabel 5
Individuele overeenstemming, betrouwbaarheid en homogeniteit van receptieve en produktieve oordelen

	Receptief					Produktief				
Stratum	F_{xy}	α	*GFI*	*Fabin*	*N*	F_{xy}	α	*GFI*	*Fabin*	*N*
0	.38	.91	1.00	.91	2101	.49	.94	1.00	.94	2202
1	.35	.91	.99	.92	2028	.48	.95	1.00	.95	2302
2	.34	.90	.99	.90	2066	.44	.93	1.00	.93	2250
3	.36	.91	.99	.91	2073	.44	.93	1.00	.93	2363

dellen en bijbehorende passingsgegevens zijn in Figuur 2 weergegeven.

3.5 Absolute verschillen

Tot nu toe is alleen aan de relatieve overeenstemming van de oordelen aandacht geschonken. De absolute overeenstemming is bepaald door een vergelijking van de gemiddelde oordelen per stratum. Uit Tabel 6 blijkt dat de leerkrachten een later beheersingsmoment aangaven naarmate zij tot een groep behoren met een hoger percentage anderstalige kleuters op



Receptief *DF*=22 *X*=291.21 *N*=2585 *GFI*=.97 *RMR*=.013
Produktief *DF*=22 *X*=173.65 *N*=2585 *GFI*=.98 *RMR*=.007

Figuur 2. Meetmodel voor schatting correlaties tussen ware stratum-scores, receptieve oordelen (parameterschattingen normaal) en produktieve oordelen (parameterschattingen cursief)

school. Leerkrachten van scholen met 70 % of meer anderstalige kleuters geven voor de receptieve beheersing als gemiddeld moment over alle woorden 0.94 aan, leerkrachten van scholen zonder anderstalige kleuters 0.71; voor de produktieve beheersing lagen deze gemiddeldes respectievelijk op 1.58 en 1.13. De verschillen tussen de groepen zijn groot te noemen. De gemiddelde receptieve en produktieve oordelen van stratum 0 en stratum 70+ verschillen respectievelijk 0.5 en 0.7 *sd.*

Tabel 6
Gemiddelde oordelen en standaarddeviaties per stratum

	Stratum			
	0	1	2	3
Receptief	.71	.88	.88	.94
	(.47)	(.48)	(.45)	(.48)
Produktief	1.13	1.47	1.52	1.58
	(.65)	(.64)	(.61)	(.60)

Het verschil tussen de extreme strata in aantallen wenselijke woorden per moment is behoorlijk groot. Volgens leerkrachten zonder anderstalige kleuters moeten bijna twee keer zoveel woorden al aan het begin van de kleutergroepen receptief beheerst worden dan volgens leerkrachten met veel anderstalige leerlingen (933 vs. 496). Voor de produktieve beheersing liggen de aantallen woorden die wenselijk geacht worden aan het begin van de kleutergroepen nog veel verder uit elkaar (430 vs. 72).

Systematische verschillen tussen gemiddelde receptieve en produktieve oordelen worden ook gevonden bij andere groepsindelingen. Leerkrachten geven een later moment aan naarmate zij een hoger percentage anderstalige kleuters in de eigen groep hebben: de verschillen tussen 0% versus 70% of meer anderstaligen in de eigen groep belopen 0.3 en 0.5 *sd.* Leerkrachten van scholen met een schoolgewicht groter of gelijk aan 1.25 noteerden systematisch latere beheersingsmomenten dan leerkrachten van scholen met een schoolgewicht kleiner dan 1.25: de effectgrootte bedraagt 0.2 respectievelijk 0.3 *sd.* Leerkrachten die alleen jongste kleuters in hun groep hadden, geven latere momenten voor de receptieve beheersing aan dan leerkrachten met alleen oudste kleuters: de effectgrootte bedraagt 0.1 *sd.* De gemiddelde produktieve oordelen verschillen niet. Ten slotte bleken ook leerkrachten met meer dan 10 jaar ervaring in het kleuteronderwijs iets latere momenten voor de receptieve beheersing aan te strepen dan leerkrachten met minder dan 10 jaar ervaring: de omvang van het effect is 0.1 *sd.* De gemiddelde produktieve oordelen verschillen opnieuw niet.

De absolute oordelen van de leerkrachten

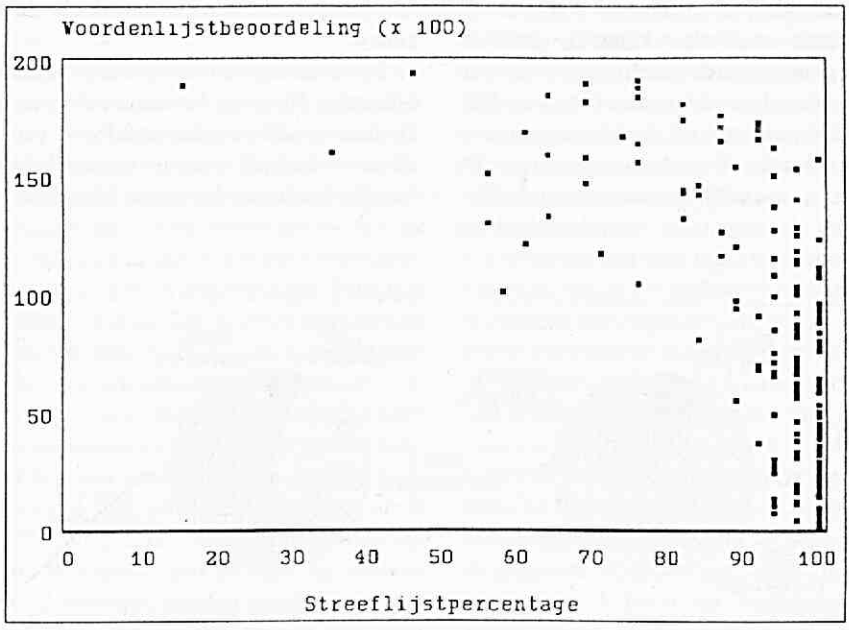
blijken afhankelijk te zijn van diverse achtergrondkenmerken van de leerkrachten, hun school en hun eigen kleutergroep. De norm van de leerkrachten versoepelt naarmate zij een hoger percentage anderstalige kleuters op school hebben, een hoger percentage anderstalige kleuters in de eigen groep hebben, verbonden zijn aan een school met een hoger schoolgewicht, jongere kleuters in de eigen groep hebben, of meer ervaring in het kleuteronderwijs hebben.

3.6 Convergentie met Nieuwe Streeflijst

De convergente validiteit van de leerkrachtoordelen is bestudeerd door na te gaan in hoeverre hun oordelen overeenstemmen met de gegevens uit de *Nieuwe Streeflijst* (Kohnstamm e.a. 1981). Uit de 2585 woorden uit de *Woordenlijst voor 4- tot 6-jarigen* (Damhuis e.a. 1992) is elk alfabetisch eerste woord met een steeds .01 hoger gemiddeld oordeel voor het moment van receptieve beheersing geselecteerd (beginnend met 0.01 en eindigend met 2.00). Van de 187 aldus gekozen woorden bleken er 39 niet of in een andere betekenis in de *Nieuwe Streeflijst* voor te komen. Voor de resterende 148 woorden is de overeenstemming tussen de Woordenlijstoordelen en de Streeflijstoordelen van Nederlandse kleuterleerkrachten bepaald.

Bij deze vergelijking moet het volgende bedacht worden. De Woordenlijstoordelen geven aan welk beheersingsmoment binnen de kleuterperiode wenselijk is. De Streeflijstoordelen geven aan of beheersing van een woord aan het einde van de kleuterperiode wenselijk is. De vraagstelling aan de beoordelaars is derhalve verschillend en we behandelen de gegevens uit de *Nieuwe Streeflijst* anders dan bedoeld, n.l. als gegevens die iets zeggen over de wenselijke verwervingsvolgorde van de woorden. Enige grond voor een dergelijke interpretatie is er overigens wel: onderzoek van Van Loon-Vervoorn (1985a en b) laat zien dat Streeflijstoordelen redelijk overeenkomen met de verwervingsleeftijd die volwassenen retrospectief rapporteren.

Verder moet bedacht worden dat de *Woordenlijst* hoofdzakelijk bestaat uit zgn. unaniemen uit de *Nieuwe Streeflijst* waardoor de verdeling van de Streeflijstoordelen zeer scheef is. Figuur 3 brengt dit in beeld voor de geselecteerde 148 woorden.



Figuur 3. Samenhang tussen oordelen in vragenlijstonderzoek en oordelen uit de Nieuwe Streeflijst

Ondanks de genoemde beperkingen is er sprake van een onmiskenbare correspondentie tussen de Streeflijstgegevens en de Woordenlijstgegevens. Uit Tabel 7 blijkt dat de woorden met het hoogste Streeflijstpercentage (nl. 100%) volgens de Woordenlijstbeoordeling duidelijk eerder beheerst moeten worden dan de woorden met een Streeflijstpercentage kleiner dan 90%. Deze overeenstemming verschaft steun aan de validiteit van de Woordenlijstgegevens.

Tabel 7
Overeenstemming tussen gegeven Nieuwe Streeflijst en Woordenlijst

Streeflijst-percentage	Receptief oordeel Woordenlijst		
	N	*Gemiddelde*	*sd*
100%	47	.54	.36
90% tot 99%	60	.75	.47
0 tot 89%	41	1.48	.34

3.7 Convergentie met feitelijke woordkennis

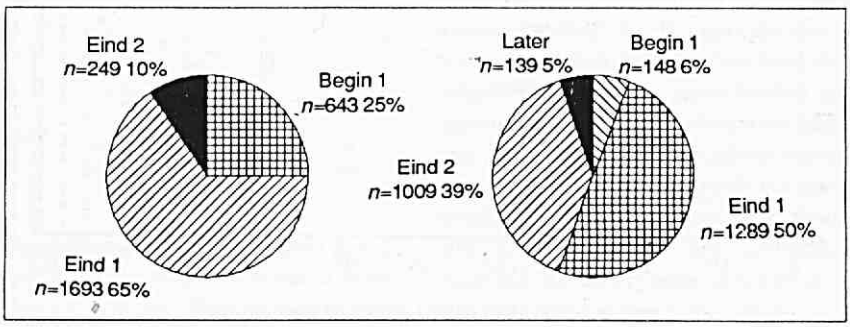
Een beperkt aantal woorden uit de *Woordenlijst voor 4- tot 6-jarigen* komt voor in de Taaltoets Allochtone Kinderen (Verhoeven, Vermeer & Van de Guchte, 1986). In het kader van een peilingsonderzoek van Boogaard e.a. (1990) is de kennis van deze woorden bij 6-jarige Nederlandse, Surinaamse en Antilliaanse, Marokkaanse en Turkse kinderen gemeten. Een vergelijking van de feitelijke woordkennis met de leerkrachtoordelen over het *wenselijke* beheersingsmoment geeft verdere steun aan de validiteit van de Woordenlijstgegevens. De *Woordenlijst* heeft 29 woorden gemeenschappelijk met de receptieve woordenschattoets van de TAK en 35 woorden met de produktieve woordenschattoets. De feitelijke woordkennis van leerlingen (ongewogen p-waarden voor de totale steekproef uit Boogaard e.a. 1990) hangt sterk samen met de leerkrachtoordelen: de pmc bedraagt voor receptieve kennis en oordelen .78 en voor produktieve kennis en oordelen .79.[3]

4 Gebruik van de Woordenlijst

4.1 Taalplan Kleuters

De gemiddelde absolute oordelen geven aan op welk moment een woord volgens de leerkrachten beheerst zou moeten worden: *het wenselijke beheersingsmoment*. Ze laten zien welke woorden kort na binnenkomst beheerst zouden moeten worden (0.01 t/m 0.50), welke aan het einde van het eerste kleuterjaar (0.51 t/m 1.50), welke aan het einde van het tweede kleuterjaar (1.51 t/m 2.00 voor receptief, 1.51 t/m 2.50 voor produktief) en welke woorden niet binnen de kleuterperiode produktief beheerst hoeven te worden (2.51 t/m 3.00). De aantallen woorden die op deze momenten beheerst zouden moeten worden volgens de oordelen van de totale groep leerkrachten, zijn afgebeeld in Figuur 4.

De verdeling van de woorden is zeer ongelijkmatig. Kort na binnenkomst zouden de kleuters al 643 woorden (een kwart van de totale woordenlijst) receptief moeten beheersen. Aan het einde van het eerste kleuterjaar komt



Figuur 4. Het aantal woorden dat volgens de gemiddelde oordelen van de leerkrachten op elk van de onderscheiden momenten beheerst zou moeten worden

daar nog eens ruim de helft van de lijst bij (1693 woorden) en voor het tweede kleuterjaar zijn er dan slechts 249 woorden over. Voor de produktieve beheersing is de verdeling van de woorden iets minder scheef. Kort na binnenkomst in de kleutergroep zouden de kleuters maar 148 woorden produktief hoeven te beheersen, aan het einde van het eerste kleuterjaar komen daar een kleine 1300 woorden bij, en voor het tweede kleuterjaar staan er nog eens 1000 woorden op het programma. Opvallend is het dat er niet veel woorden zijn die de kleuters nog niet produktief hoeven te beheersen aan het einde van de kleuterperiode (139). Tot deze groep behoren moeilijker telwoorden (16 t/m 19, de tientallen vanaf 30, en 100 en 1000), hoeveelheidswoorden (zoals 'meter', 'liter', 'kilo', 'minuut', 'kwart'), de namen van de maanden, schoolse woorden uit hogere groepen (bijvoorbeeld 'huiswerk', 'invullen', 'rekenles', 'spreekbeurt'), tamelijk abstracte begrippen of relaties (zoals 'overmorgen', 'laatst (onlangs)', 'behalve', 'trouwens'), enkele zuidelijke varianten van woorden (zoals 'gij', 'opdekken'), enkele woorden die voor kleuters niet zo relevant zijn (zie par. 5) en namen van landen en daarvan afgeleide bijvoeglijk naamwoorden (waaronder 'Turkije', 'Spaans', 'De Antillen').

Gezien de systematische verschillen in absolute oordelen tussen de groepen leerkrachten (par. 3.5), zouden voor de toepassing van de gegevens in Taalplan Kleuters de absolute oordelen van die groep leerkrachten gebruikt kunnen worden, die het meest overeenkomt met de doelgroep van Taalplan Kleuters: kleutergroepen met 70 % of meer anderstaligen. Maar ook de absolute oordelen van de groep leerkrachten van scholen met 70 % of meer anderstalige kleuters vertonen een sterk onevenredige verdeling van de woorden over de kleuterperiode: kort na binnenkomst zouden 496 woorden receptief beheerst moeten worden, aan het einde van groep 1 nog eens 1694 woorden, en in groep 2 zouden er nog maar 395 woorden bijkomen. Produktief zouden kort na binnenkomst 72 woorden moeten worden beheerst, aan het einde van groep 1 komen daar 1054 woorden bij, aan het einde van groep 2 nog eens 1250; van 209 woorden wordt de produktieve beheersing niet binnen de kleuterperiode noodzakelijk geacht.

Voor het Taalplan Kleuters was een gelijkmatiger verdeling van de woorden wenselijk. Het lesmateriaal van Taalplan Kleuters zou regelmatig gespreid over de twee kleuterjaren gebruikt moeten gaan worden: in 7 taalactiviteiten per week, met ongeveer 6 nieuwe woorden per activiteit. Om deze gelijkmatige verdeling te krijgen is de rangordening van de woorden als relatief gegeven gebruikt. Over de rangordening van de woorden bestaat tussen groepen leerkrachten vrijwel perfecte overeenstemming (zie par. 3.4).

Voor Taalplan Kleuters is op grond van een eerste reeks proefinvoeringsgegevens vastgesteld hoeveel woorden in totaal aangeboden zouden kunnen worden. Dit zouden er ruim 2000 kunnen zijn[4]. Ongeveer 400 woorden die volgens de rangordening het laatst beheerst zouden moeten worden, zijn niet in het Taalplan opgenomen. De 2000 woorden die eerder in de *Rangordelijst receptieve beheersing* (zie par. 4.2) staan, zijn in twee groepen verdeeld: de eerste groep van circa 1000 woorden is opgenomen in het lesmateriaal voor het eerste kleuterjaar, de tweede groep van 1000 woorden in het materiaal voor het tweede kleuterjaar. Binnen elk kleuterjaar is niet vastgehouden aan de rangordening, omdat het lesmateriaal rond thema's is opgezet en de woorden binnen thema's in betekenisvolle contexten aangeboden moeten worden. Na een tweede proefinvoeringsjaar, bleek het aantal woorden verder teruggebracht te moeten worden, tot ongeveer 800 woorden per kleuterjaar.

4.2 Woordenlijst voor 4- tot 6-jarigen

Op basis van de gemiddelde oordelen over het receptieve beheersingsmoment van de totale groep leerkrachten is een *Rangordelijst receptieve beheersing* opgesteld. Deze lijst is de kern van *Woordenlijst voor 4- tot 6-jarigen* (Damhuis e.a. 1992). De lijst is ook in alfabetische volgorde van de woorden opgenomen, de *Alfabetische lijst*. Deze beide lijsten staan in Deel II, voorafgegaan door een uitleg bij de lijsten en de gegevens die daarin per woord zijn opgenomen. Deel I bevat de toelichting bij de *Woordenlijst* en het vragenlijstonderzoek.

Woordenlijst voor 4- tot 6-jarigen is een minimumstreeflijst woordenschat voor jonge kinderen, ongeacht hun herkomst (zie par. 2.3). De *Rangordelijst receptieve beheersing* kan als

uitgangspunt dienen wanneer men de relatieve volgorde van woorden wil gebruiken, of woorden zoekt die bij een bepaald moment of een bepaalde periode binnen de kleutertijd passen. De *Alfabetische lijst* is bestemd voor het nazoeken van gegeven oordelen over bepaalde woorden.

Afhankelijk van de specifieke gebruikerswensen en -omstandigheden kunnen de absolute oordelen of de rangordening richtinggevend zijn. Wanneer de absolute oordelen gebruikt worden, moet men zich rekenschap geven van de verschillen tussen leerkrachten uit verschillende strata. Het kan wenselijk zijn om niet de oordelen van de totale groep te gebruiken, maar de oordelen van leerkrachten van scholen die qua percentage anderstalige kleuters overeenkomen met de scholen of onderwijssituaties waarvoor men de lijst wil gebruiken. De *Woordenlijst* bevat daartoe een omrekeningstabel (Damhuis e.a. 1992, p.53). Vooral voor scholen zonder anderstalige kleuters en voor scholen met 70 % of meer anderstalige kleuters is deze omrekening zinvol: voor deze groepen wijken de oordelen het meest af van het totaaloordeel (zie ook par. 3.5).

De *Woordenlijst voor 4- tot 6-jarigen* kan zowel in het reguliere onderwijs Nederlands als in het onderwijs Nederlands als tweede taal als leidraad gehanteerd worden. De lijst zou gebruikt kunnen worden bij de invulling van het woordenschatonderwijs voor jonge kinderen, ongeacht hun herkomst. Ontwikkelaars van les- en toetsmateriaal, schoolbegeleiders en leerkrachten zouden voor een bepaald onderwijsmoment woorden uit de lijst kunnen selecteren (gemiddeld moment), of een rangordening aanbrengen in aan te bieden woorden voor een bepaalde periode (onderlinge rangorde van de woorden). Zo zou er gebruik gemaakt kunnen worden van de lijst in geïntegreerd onderwijs Nederlands als tweede taal in kleutergroepen, in intensieve opvangprogramma's Nederlands als tweede taal voor jonge neveninstromers (zowel in centrale opvang als in andere modellen, zie Blok, Emmelot & De Kat, 1992), in reguliere kleutergroepen voor het taalonderwijs Nederlands als moedertaal, en, ten slotte, ook in de begeleiding van jonge kinderen in peuterspeelzalen en kinderdagverblijven (met name de woorden waarvan de beheersing al kort na binnenkomst in de kleutergroepen wenselijk geacht wordt; zie Evegroen & Roselaar, 1992).

5 Discussie

Bij de *Woordenlijst 4- tot 6-jarigen* kunnen verschillende vragen gesteld worden. Een eerste vraag is of de Woordenlijst wel de juiste, of beter gezegd de meest relevante woorden bevat. De kwaliteit van de lijst hangt in hoge mate samen met de kwaliteit van de bronlijsten van Kohnstamm e.a. (1981) en Coenen en Vermeer (1988). Tegen de uit de eerste lijst gekozen unaniemen, 1838 in getal, lijkt ons weinig bezwaar te kunnen bestaan. Aan de lijst van Coenen en Vermeer kleven beperkingen die in paragraaf 2.1. reeds aangegeven zijn. De wijze waarop wij deze lijst gebruikt hebben, brengt met zich mee, dat er ook woorden zijn opgenomen die een evident lage relevantie voor kleuters hebben (bijvoorbeeld 'whisky', 'sexy', 'flipperkast'). Daartegenover staan andere laag-frequente woorden die wel duidelijk relevant zijn (bijvoorbeeld 'rietje', 'kapstok', 'vakantie', 'vastmaken', 'schaar', 'opdrinken'). De laag-frequente irrelevante woorden konden door de beoordelaars met een 'laat' wenselijk beheersingsmoment beoordeeld worden. Voor woorden als 'whisky' en 'flipperkast' gaven de leerkrachten inderdaad aan dat ze pas laat in de kleutergroepen (receptief) of niet in de kleutergroepen (produktief) beheerst hoefden te worden. De beheersing van wèl relevante laagfrequente woorden als 'schaar' en 'opdrinken' wordt al heel vroeg wenselijk geacht[5]. Tegen de thema-gebonden woorden uit het materiaal van Taalplan Kleuters kan, vanuit een gezichtspunt dat deze specifieke methode overstijgt, vanzelfsprekend bezwaar gemaakt worden. Ondanks deze beperkingen lijkt de *Woordenlijst 4- tot 6-jarigen* voornamelijk welrelevante woorden te bevatten. Het aandeel van de unaniemen in de Woordenlijst is groot. Bovendien is de overlap tussen Kohnstamm e.a. (1988) en Coenen en Vermeer (1988) aanzienlijk: de gemaakte selectie uit Coenen en Vermeer bevatte al 1572 van de unaniemen. Het valt natuurlijk niet uit te sluiten dat belangrijke woorden in de woordenlijst ontbreken. Nader onderzoek, bijvoorbeeld naar de mate waarin de lijst dekkend is voor het woordgebruik in

kleutergroepen, zou lacunes kunnen uitwijzen.

Een tweede vraag is of de lijst niet een te lage doelstelling voor de woordenschat vormt? Bevat de lijst niet teveel woorden die door kleuters al beheerst worden? Aan de hand van gegevens van Boogaard e.a. (1990) kan deze vraag ontkennend beantwoord worden, zie Tabel 8. De 29 woorden die de Woordenlijst gemeenschappelijk heeft met de receptieve woordenschattoets uit de TAK (Verhoeven, Vermeer & Van de Guchte, 1986) worden aan het einde van groep 2 door Nederlandstalige leerlingen goeddeels beheerst: 25 van de 29 woorden hebben een *p*-waarde van boven .80. Voor de anderstalige leerlingen geldt dit zeker niet. De Turkse leerlingen beheersen maar de helft van de 29 woorden. Beschouwing van de 35 woorden die in de produktieve woordenschattoets van de TAK voorkomen laat zien dat de produktieve beheersing nog verder bij de doelstelling achterblijft. Nederlandstalige leerlingen beheersen 22 van de 35 woorden; Turkse leerlingen slechts 5 van de 35.

Een volgende vraag is of de lijst voldoende 'stof' bevat voor twee kleuterjaren woordenschatonderwijs? Volgens de absolute oordelen van de leerkrachten zou een fors deel van de woorden al aan het einde van het eerste kleuterjaar beheerst moeten worden (zie par. 4.1). Uit de proefinvoeringsgegevens van Taalplan Kleuters blijkt dit echter binnen de huidige mogelijkheden van het reguliere onderwijs in kleutergroepen geenszins haalbaar te zijn. Hoewel kinderen ook via incidenteel leren hun woordenschat uitbreiden, kan de verwerving van woorden die tot het streefdoel van het onderwijs behoren, o.i. niet overgelaten worden aan het incidentele, ongestuurde leren. De streefwoorden moeten juist onderdeel gemaakt worden van intentioneel leren door ze op te nemen in lesprogramma's. In de onderwijspraktijk is er een limiet aan het maximaal aantal woorden dat middels intentioneel leren aangeboden kan worden. In de literatuur worden maxima tussen de 5 en 10 woorden per les van 20 tot 30 minuten genoemd (Coenen & Vermeer, 1988, p. 17). Bij de proefinvoering van Taalplan Kleuters is gebleken dat er in de normale kleutergroeppraktijk (d.w.z. zónder extra voorzieningen) maximaal 6 activiteiten per week aan gerichte taalstimulering besteed kunnen worden. In een schooljaar kunnen ongeveer 30 weken effectief aan een lesprogramma besteed worden. De onder de huidige omstandigheden maximaal haalbare intentionele uitbreiding van de woordenschat zou hiermee op 6 x 6 x 30 = 1080 woorden per jaar komen. Maar een deel van de activiteiten moet ook besteed worden aan het voor woordenschatverwerving zo belangrijke herhalen van woorden. In het definitieve materiaal van Taalplan Kleuters zullen daarom circa 800 woorden per kleuterjaar aangeboden worden. Hoewel dit aantal momenteel praktisch het hoogst haalbare is, ligt het dus ver onder het eigenlijk wènselijke aantal.

Ter vergelijking van de omvang van het woordaanbod in Taalplan Kleuters kan verwezen worden naar het woordenschatexperiment van Appel en Vermeer (1992; 1993). In het experiment worden ruim 1000 woorden per jaar aangeboden. Vanwege tijdgebrek worden echter op lang niet alle scholen alle activiteiten uitgevoerd.

De conclusie moet zijn dat de *Woordenlijst* voldoende stof voor twee jaar kleuteronderwijs bevat. Zelfs met een vrij intensief taalstimuleringsmodel als Taalplan Kleuters kan de reguliere kleuterleerkracht de eigenlijk wenselijke woordenschatontwikkeling niet bewerkstelligen. Dit roept de vraag op of er niet méér mogelijkheden voor taalstimulering geschapen moeten worden.

Tabel 8

Aantallen woorden die feitelijk beheerst worden aan het einde van de kleuterperiode (Receptief = in Passieve Woordenschattaak (PW); Produktief = in Actieve Woordenschattaak (AW); Beheerst = p > .80; Niet-beheerst = p ≤ .80)

	Receptief (PW)		Produktief (AW)	
	Beheerst	Niet-beheerst	Beheerst	Niet-beheerst
Nederlands	25	4	21	14
Surinaams/Antilliaans	16	13	15	20
Marokkaans	15	14	6	29
Turks	14	15	5	30

Wat, ten slotte, is de betekenis van de geconstateerde verschillen tussen leerkrachten van scholen zonder anderstalige leerlingen en leerkrachten van scholen met erg veel anderstalige leerlingen? De duidelijk verschillende doelstellingen vormen een aanwijzing voor een lager streefniveau op scholen met veel anderstalige leerlingen. Het verschil moet ons inziens niet opgevat worden als een bewijs van verondersteld slechter onderwijs op zogenoemde *zwarte* scholen. Het verschil weerspiegelt de grote verschillen in woordkennis tussen Nederlandstalige en anderstalige leerlingen en lijkt een afspiegeling van wat onder de huidige omstandigheden in kleutergroepen met veel anderstaligen feitelijk haalbaar is. De verschillen in streefniveaus benadrukken, wellicht ten overvloede, de noodzaak van versterkte aandacht voor de woordenschatontwikkeling, en meer in het algemeen, voor de ontwikkeling van het Nederlands als tweede taal van anderstaligen.

Noten

1 Het ontbreken van de antwoordmogelijkheid 'Later' bij de receptieve oordelen kan geleid hebben tot een relatief groter aantal ontbrekende waarnemingen bij woorden die men eigenlijk in deze ontbrekende categorie had willen rangschikken. Leerkrachten kunnen voor zulke woorden óf de laatst mogelijke categorie hebben gekozen, óf inderdaad geen oordeel genoteerd hebben. Dat laatste is nogal eens het geval, blijkens de correlatie tussen het aantal ontbrekende oordelen en het over de leerkrachten gemiddelde wenselijke beheersingsmoment (r=.56). Het effect van dit alles op de beoordeling is overigens miniem. Als bij de 41 woorden waarvoor 7 of meer oordelen ontbreken (i.e. minstens 10% ontbrekende oordelen) de ontbrekende oordelen vervangen worden door het oordeel 'einde groep 2', verschuift het gemiddelde van deze groep woorden van 1.84 naar 1.85.

Het is verder overigens de vraag of er wel veel woorden wat betreft de wenselijke receptieve beheersing in de categorie 'later dan de kleutergroepen' zouden zijn geplaatst door de leerkrachten, als ze wèl die keuze hadden gehad. Bij de beoordeling van de produktieve beheersing zijn slechts 139 woorden gemiddeld genomen in de categorie 'Later' terecht gekomen. Voor de receptieve beheersing zou zeker een lager getal gevonden zijn, daar de leerkrachten het wenselijke moment van receptieve beheersing beduidend eerder leggen dan het produktieve beheersingsmoment.

2 De onder het model geschatte correlaties tussen de ware jury-scores zijn overigens vrijwel precies gelijk aan de voor attenuatie gecorrigeerde geobserveerde jury-scores.

3 Met betrekking tot dezelfde 29 woorden vertonen de streeflijstoordelen een correlatie van .63 met de feitelijke woordkennis.

4 Voor een meer gedetailleerde beschrijving van het gebruik van de resultaten van het vragenlijstonderzoek in Taalplan Kleuters, raadplege men hoofdstuk 4 van *Woordenlijst voor 4- tot 6- jarigen* (Damhuis e.a., 1992). Daarin wordt onder meer de rol van woorden uit twee programma's voor kleuters (Van Kleef, 1988; Boers & Van Kleef, 1990) en van 'Algemene woorden' en 'Themawoorden' besproken.

5 Voor de hier genoemde laag-frequente woorden uit de lijsten van Coenen en Vermeer (1988) zijn de leerkrachtoordelen over het wenselijke beheersingsmoment als volgt (Rec. = Receptief; Prod. = Produktief):

Relevante woorden			Niet-relevante woorden		
	Gemiddeld wenselijk moment			Gemiddeld wenselijk moment	
Woord	Rec.	Prod.	Woord	Rec.	Prod.
schaar	0.08	0.37	sexy	1.77	2.63
rietje	0.20	0.54	flipperkast	1.83	2.77
opdrinken	0.20	0.66	whisky	1.94	2.88
kapstok	0.31	0.70	adelaar	1.95	2.82
vakantie	0.48	0.92			
vastmaken	0.52	1.07			

Literatuur

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23*(3), 285-303.

Appel, R., & Vermeer, A. (1992). Duizend woorden per jaar. Rapportage NT2-experiment in de onderbouw. *Stimulans, 10*(7), 11-13.

Appel, R., & Vermeer, A. (1993). Woordenschat van allochtone leerlingen met sprongen vooruit. *Didaktief, 23*(3), 5-7.

Blok, H., Emmelot, Y., & Kat, E. de (1992). *De opvang van neveninstromers in de basisschool; een eerste verkenning. SCO-rapport 288.* Amsterdam: Stichting Centrum voor Onderwijsonderzoek.

Boers, M., & Kienstra, M. (1993). *Taalplan Kleuters: Werkplannen groep 2.* Rotterdam: Projectbureau OVB Rotterdam.

Boers, M., Kienstra, M., & Padmos, M. (1993). *Taalplan Kleuters: Werkplannen groep 1.* Rotterdam: Projectbureau OVB Rotterdam.

Boers, M., & Kleef, M. van (1990). *Laat wat van je horen. Nederlandse taal voor kleuters in meertalige groepen (deel 2).* Rotterdam: Projectbureau OVB Rotterdam.

Boogaard, M., Damhuis, R., Glopper, K. de, & Bergh, H. van den (1990). *De Nederlandse taalvaardigheid van allochtone en Nederlandse kleuters. Peiling van de taalvaardigheid van Surinaamse, Antilliaanse, Marokkaanse, Turkse en Nederlandse leerlingen aan het einde van de kleuterperiode* (Forum 4). Amsterdam/Lisse: Swets & Zeitlinger.

Coenen, M., & Vermeer, A. (1988). *Nederlandse woordenschat allochtone kinderen.* Tilburg: Zwijsen.

Damhuis, R., Glopper, K. de, Boers, M., & Kienstra, M. (1992). *Woordenlijst voor 4- tot 6-jarigen: een streeflijst voor kleuters.* Rotterdam: Projectbureau OVB Rotterdam.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. MCKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition,* (pp.73-87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Evegroen, R., & Roselaar, T. (1992). Meertaligheid in kindercentra. Een nieuw initiatief van het Vier Steden Project Welzijn. *Stimulans, 10*(5), 12-13.

Glopper, K. de, Daalen-Kapteijns, M. van, & Schouten-van Parreren, C. (1992). *Woordenschatverwerving in moedertaal en vreemde talen: een kwestie van transfer?* Paper gepresenteerd op de conferentie voor Het Schoolvak Nederlands, Antwerpen, 17 oktober 1992.

Glopper, K. de, Daalen-Kapteijns, M. van, & Schouten-van Parreren, C. (1993). *Predictors of vocabulary learning in mother tongue and foreign languages.* Paper presented at the Xth World Congress of Applied Linguistics, August 8-14, Amsterdam.

Hagglund, G. (1982). Factor analysis by instrumental variables methods. *Psychometrika, 47,* 209-222.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis.* New York: John Wiley & Sons.

Jenkins, J. R., Matlock, B., & Slocum, T. A. (1989). Two approaches to vocabulary instruction: The teaching of individual word meanings and practice in deriving word meaning from context. *Reading Research Quarterly, 24*(2), 253-279.

Kleef, M. van (1988). *Knoop het in je oren. Nederlandse taal voor kleuters in meertalige groepen.* Rotterdam: Projectbureau OVB Rotterdam.

Kohnstamm, G. A., Schaerlaekens, A. M., Vries, A. K. de, Akkerhuis, G. W., & Froonincksx, M. (1981). *Nieuwe Streeflijst Woordenschat voor 6-jarigen gebaseerd op onderzoek in Nederland en België.* Lisse: Swets & Zeitlinger.

Loon-Vervoorn, W. A. van (1985a). De Nieuwe Streeflijst als maat voor de verwervingsleeftijd van woorden. *Nederlands Tijdschrift voor de Psychologie, 40,* 44-47.

Loon-Vervoorn, W. A. van (1985b). De Nieuwe Streeflijst als maat voor de verwervingsleeftijd van woorden II. *Nederlands Tijdschrift voor de Psychologie, 40,* 503-506.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304-330.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. MCKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp.19-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schouten-van Parreren, C., Glopper, K. de, & Daalen-Kapteijns, M. van (1993). *Vocabulary and metacognition in mother tongue and foreign languages.* Paper presented at the Xth World Congress of Applied Linguistics, August 8-14, Amsterdam.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. MCKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp.89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vallen, T., & Stijnen, S. (1991). Ethnic minorities and ethnic minority languages in Dutch primary schools: problems and challenges. *Language and Education, 5,* 113-123.

Verhallen, M. (1991). *Woordenschatuitbreiding bij anderstalige kinderen. ATW-publikatie nr. 57.* Amsterdam: Instituut voor Algemene Taalwetenschap.

Verhoeven, L., & Vermeer, A. (1985). Ethnic group differences in children's oral proficiency in Dutch. In G. Extra & T. Vallen (Eds.), *Ethnic minorities and Dutch as a second language* (pp.105-131). Dordrecht: Foris.

Verhoeven, L., & Vermeer, A. (1992). Woordenschat van leerlingen in het Basis- en MLK-onderwijs. *Pedagogische Studiën, 69,* 218-234.

Verhoeven, L., Vermeer, A. & Guchte, C. van de (1986). *Taaltoets voor allochtone kinderen.* Tilburg: Zwijsen.

Manuscript aanvaard 19-4-1993

Auteurs

R. Damhuis is als onderwijsonderzoeker werkzaam aan het SCO-Kohnstamm Instituut van de Universiteit van Amsterdam.

K. de Glopper is hoogleraar in de pedagogische en onderwijskundige wetenschappen aan de Universiteit van Amsterdam.

Adres: SCO-Kohnstamm Instituut, Grote Bickerstraat 72, 1013 KS Amsterdam

Abstract

Teachers' judgements of the desirable vocabulary for 4- to 6-year old children: reliability, validity and possibilities for application.

R. Damhuis & K. de Glopper. Pedagogische Studiën, 1993, 70, 322-336.

Connected with the development of *Taalplan Kleuters*, a model for the stimulation of Dutch language development in 4- to 6-year old multilingual children, a questionnaire survey was conducted in which teachers judged the desirable proficiency moment of 2585 words constituting the basis of Taalplan Kleuters. The survey was meant to obtain an ordering of these words within the pre-school period. This article discusses the reliability and the validity of the judgements, and the possibilities for the application of the resulting list of classified words. The absolute judgements differed according to the percentage of multilingual children in the schools. With respect to the relative judgements the teacher groups agreed strongly. Teacher judgements of the desirable proficiency moments were closely related to students' actual word knowledge. The list appears to be not too low a target for vocabulary education for 4- to 6-year old children.

Spelen op school

E.G. Harskamp

Samenvatting

Zes jaar na het van kracht worden van de Wet op het Basisonderwijs is de discussie in volle gang over de kwaliteit van het kleuteronderwijs. Velen zijn bang dat het spel als een belangrijke verworvenheid van het kleuteronderwijs verloren gaat op de huidige basisschool. Het onderzoek laat zien dat in groep 2 van de basisschool grote verschillen zijn in de mate waarin leerkrachten kinderen laten spelen en het spel van kleuters stimuleren. Over het algemeen wordt ongeveer een derde deel van de schooltijd gebruikt voor spel en de rest van de tijd voor andere activiteiten. De verschillen in de mate van spelstimulering tussen leerkrachten hangen samen met de mate van spelbetrokkenheid van de kleuterklassen. Meer dan spelstimulering blijkt echter de mate waarin de leerkrachten een efficiënte klasse-organisatie hebben, samen te hangen met de spelbetrokkenheid van de kinderen. De praktische aanbevelingen vanuit dit project hebben betrekking op het verder uitwerken van spelprogramma's voor het kleuteronderwijs en de inservice training van kleuterleerkrachten in klassemanagementvaardigheden.

Inleiding

In 1985 zijn de kleuterschool en de lagere school formeel samengevoegd tot één basisschool voor vier- tot twaalfjarigen. In de basisschool dient een doorgaande lijn in het onderwijs aan kleuters en lagere school kinderen te worden aangeboden. In de periode na de totstandkoming van de basisschool hebben vele kleuterdeskundigen hun verontrusting geuit over de verschraling van het onderwijs aan kleuters en het verlies van de verworvenheden van het vroegere kleuteronderwijs. Leerkrachten in groep een en twee van de basisschool zouden minder aandacht besteden aan kleuters met ontwikkelingsproblemen en minder gelegenheid geven voor spel en vrije exploratie dan in het kleuteronderwijs gebruikelijk was (zie Janssen-Vos, 1991). Er zijn enige aanwijzingen die de verontrusting kunnen staven. Ervaren kleuterleidsters verlaten de basisschool en worden vervangen door leerkrachten die geen specifieke opleiding voor onderwijs aan vier tot zesjarigen hebben (zie Van de Ven, Vrieze & Claessen, 1988). De opleiding van leerkrachten van het lager onderwijs en voor het kleuteronderwijs zijn geïntegreerd. In de nieuwe lerarenopleiding voor het basisonderwijs (PABO) zijn er bijna geen specialisten op het gebied van het jonge kind en de opleidingen bereiden leerkrachten niet of nauwelijks voor op het onderwijs aan kleuters zoals dat op de vroegere opleiding voor kleuterleidsters (KLOS) gebeurde (zie HBO-raad, 1989; Van der Duim & Blok, 1990). Inmiddels is door de minister van onderwijs besloten dat de PABO een afstudeerspecialisatie over het jonge kind krijgt.

Een stroom van nieuwe leermaterialen op het gebied van voorbereidend lezen, schrijven en rekenen leggen een druk op de leerkrachten in groep een en twee om vroegtijdig met een gestructureerd aanbod aan leerstof te beginnen.

De verontrusting van kleuterdeskundigen heeft in onderwijsvakbladen en in de landelijke pers de aandacht getrokken.

Voor het Ministerie van Onderwijs en Wetenschappen was de ongerustheid over kwaliteit van het onderwijs aan jonge kinderen aanleiding om een onderzoek te laten doen naar het spel op de basisschool.

Het onderzoek dient zich te richten op de vraag wat leerkrachten doen om het spel van kleuters te stimuleren en welk effect spelstimulering heeft op het spelgedrag van kinderen.

Vanuit de traditie van het kleuteronderwijs wordt verondersteld dat de cognitieve en sociale vaardigheden van kinderen zich meer ontwikkelen door kinderen gelegenheid te geven tot spel en vrije exploratie dan door het aanbieden van gestructureerde onderwijspro-

gramma's. In Nederland is hierover reeds geruime tijd een discussie gaande (zie Van Kuyk, 1987). De literatuurstudie van Christie en Johnson (1983) over de cognitieve en sociale effecten van experimenten met spelinterventies in Angelsaksische landen en Levy's literatuurstudie van spelinterventies en het effect op de taalontwikkeling van kinderen (Levy, 1984), lijken de hiervoor genoemde assumptie te ondersteunen. Ook in Nederlands onderzoek naar spelinterventies worden positieve effecten van spelinterventies op cognitieve en sociale vaardigheden van kinderen in de basisschool gerapporteerd (zie Van der Kooy & Rost, 1985; Hellendoorn & Van Zanten, 1989).

Vele onderzoekers wijzen echter op de methodologische problemen in het effectonderzoek van spelinterventies. Er zijn vaak geen adequate alternatieve treatments in de experimenten of het onderscheid tussen spelinterventies en andere interventies (bijvoorbeeld vrije exploratie of directe instructie) is niet goed geoperationaliseerd. Deze problemen maken een definitieve conclusie over de waarde van spelstimulering voor de ontwikkeling van jonge kinderen voorbarig.

Spelstimulering en andere factoren die het spelgedrag van kinderen kunnen beïnvloeden

De centrale onderzoeksvraag is: "Wat doen leerkrachten om het spel van kleuters te stimuleren en met welk effect voor het spelgedrag?". Om deze vraag te beantwoorden worden factoren besproken die het spelgedrag van kleuters in de schoolsituatie zullen beïnvloeden. Allereerst wordt uiteengezet hoe een passende definitie is gekozen van het begrip spelgedrag.

Spelgedrag

In de onderzoeksliteratuur over spel staat vrijwel steeds de vraag centraal: 'Wat verstaan we onder spel?' (zie Rubin, Fein & Vandenberg, 1983). Om spel te onderscheiden van andere activiteiten van kinderen, zoals exploreren, passief kijken of ruziën wordt aan spelgedrag vaak de algemene typering gegeven:

- spel heeft vaak een 'doen alsof' karakter waarbij het kind symbolisch handelt en zich cognitieve structuren eigen maakt;
- bij spelen is een kind vrijwillig bezig en intrinsiek gemotiveerd;
- spelen is doel op zich, het gaat kinderen niet zozeer om een eindprodukt maar om het spel als ontspannende of spannende activiteit;
- bij spelen is een kind actief bezig waarbij het probeert zichzelf en de spelsituatie onder controle te krijgen.

Er kleven veel bezwaren aan het geven van een sluitende definitie van spelgedrag. Afhankelijk van de theoretische benadering die men kiest, worden bepaalde kenmerken geaccentueerd. Zo kan men spel beschouwen vanuit de motorische functie van spel (zie o.a. Bühler, 1928), de cognitieve functie van spel (zie o.a. Piaget, 1962), de sociale en emotionele functie van spel (zie o.a. Erikson, 1974), of de motivationele functie van spel (zie Berlyne, 1960). Vanuit elk van deze invalshoeken komt men tot het beklemtonen van andere kenmerken van spelgedrag.

Zo zal in het sensomotorisch spel als touwtje springen of tikkertje vaak geen sprake zijn van 'doen alsof', terwijl in het symbolisch spel in de poppenhoek hiervan wel sprake is. Wanneer kleuters een bouwwerk van blokken maken (constructiespel), streven ze vaak nadrukkelijk naar een eindprodukt. Kortom, enkele van bovengenoemde kenmerken zijn niet op alle spelvormen van toepassing. De meeste onderzoekers van spelgedrag van kinderen hanteren dan ook geen allesomvattende definitie van spelgedrag maar beperken zich in hun onderzoek tot één of enkele spelvormen (zie Christie & Johnson, 1983). In onderzoek naar spel en spelstimulering op de basisschool moet echter een algemene definitie van spelgedrag worden gekozen. In kleuterklassen worden immers meerdere spelvormen naast en door elkaar gehanteerd (zie Janssen-Vos, Dikken & Verkley, 1990).

Hoewel spelgedrag niet zonder meer te onderscheiden is van ander gedrag, zijn in het kleuteronderwijs de situaties waarin spelgedrag wordt verwacht vrij nauwkeurig aan te geven. Het betreft tafels en hoeken waar spelmaterialen zijn opgesteld en waarbinnen kinderen over het algemeen vrij kunnen kiezen welke activiteiten ze willen doen. Bosch en Bruins (1984) geven een overzicht van spelsituaties in

de klas, zoals de poppenhoek, de zandtafel, het winkeltje of de bouwhoek en de constructietafel. Spelsituaties verschillen van werksituaties (zoals puzzels maken, knippen en plakken of het maken van werkbladen met voorbereidende lees- of schrijfoefeningen) omdat bij deze laatste de leerkracht een opdracht geeft of dat het materiaal de mogelijkheden voorschrijft en kinderen weinig vrijheid hebben om zelf het doel van hun activiteiten te bepalen.

Er zijn verschillende invalshoeken om het spelgedrag van kinderen te beschrijven (zie Harskamp, Pijl & Snippe, 1991). In dit onderzoek waarin het spelaanbod van de leerkracht centraal staat en niet het spelgedrag van individuele leerlingen wordt een definitie van spelgedrag gekozen die op allerlei spelvormen en spelsituaties van toepassing is en die eenduidig kan worden gehanteerd.

De meest essentiële kenmerken van spelgedrag zijn vanuit de onderwijspraktijk bezien, de vrijheid van keuze en de intrinsiek gemotiveerde en gecontroleerde activiteit van een kind gedurende het verblijf in een spelsituatie.

Een indicatie van gemotiveerde en gecontroleerde activiteit van kinderen in spelsituaties is de aangehouden en gerichte aandacht van kinderen (zie Ruff & Lawson, 1990). Aangehouden en gerichte aandacht gedurende een bepaalde periode zijn uitingen van informatieverwerkingsvaardigheden waarbij motivatie en zelfcontrole ten aanzien van selectie en verwerking van informatie een grote rol spelen.

Ruff en Lawson wijzen op onderzoeksresultaten over de relatie tussen aangehouden en gerichte aandacht en inhoudelijke complexiteit van het spelgedrag (zie ook Kopp & Vaughn, 1982). Laevers (1991) beklemtoont het belang van niveaus van betrokkenheid van kinderen in kleuterklassen. Bij betrokkenheid speelt de aangehouden en gerichte aandacht een grote rol. De betrokkenheid wordt gedurende een korte periode geobserveerd en beoordeeld ten aanzien van de duur en gerichtheid van de aandacht van het kind. Laevers en zijn medewerkers bezien gerichte en aangehouden aandacht niet zozeer als een maat voor zelfcontrole in het selecteren en verwerken van informatie, maar als een maat voor doelgericht en intrinsiek gemotiveerd bezig zijn met activiteiten in spelsituaties.

In dit onderzoek wordt spelgedrag geoperationaliseerd als 'het niveau van betrokkenheid dat kinderen tonen in spelsituaties in de klas'. Er worden vier niveaus van betrokkenheid onderscheiden n.a.v. Laevers (1991):

1. geen spel (het kind kijkt rond, loopt weg van de spelsituatie etc.);
2. oppervlakkig spel (het kind kijkt de helft van de tijd ergens anders naar dan naar het spel, doelloos gedrag zoals, het continu heen en weer schuiven van een autootje, over iets anders praten gedurende de helft van de tijd);
3. tamelijk intensief spel (gerichte aandacht gedurende het grootste gedeelte van de geobserveerde tijd);
4. zeer intensief spel (gerichte aandacht gedurende vrijwel de gehele geobserveerde tijd).

Spelstimulering

Verschillende factoren zullen het spelgdrag (c.q. de mate van betrokkenheid van kinderen in spelsituaties) beïnvloeden. Vanuit de vraagstelling van het onderzoek is het vooral van belang om na te gaan of op spelstimulering gerichte gedragingen van leerkrachten tot het beoogde effect leiden.

Er zijn verschillende typen van spelinterventies (bemoeienissen van leerkracht of leidster met het spelgedrag van kinderen). In sociodramatische spel-experimenten (zie Singer, 1985; Hellendoorn & Van Zanten, 1989; Faber, 1990) hebben onderzoekers een aantal verschillende spelinterventies ontwikkeld en vormgegeven in draaiboeken. Meestal gaat het om het stimuleren van symbolisch spel van kleuters met ontwikkelingsproblemen. In de spelsessies introduceert de spelleider een spelthema met behulp van passend materiaal en begint zelf te spelen terwijl ze kinderen uitnodigt om mee te doen. In latere spelsessies trekt de spelleider zich meer terug na de introductie en worden kinderen uitgenodigd het initiatief over te nemen. Vaak is er een groot aantal spelthema's, oplopend in complexiteit, dat achtereenvolgens wordt aangeboden tijdens de sessies. Bovengenoemde onderzoekers rapporteren positieve invloeden van spelstimulering op het spelgedrag van kinderen.

In het onderwijs aan kleuters zal spelstimulering vaak een meer algemeen motiverend en minder sturend karakter hebben dan in de

experimenten die hierboven zijn aangeduid. De mate waarin de leerkracht algemeen stimulerend dan wel sturend optreedt in spelsituaties zal mede afhankelijk zijn van de opvattingen die zij heeft over het belang van gerichte spelsuggesties voor kinderen die moeilijk tot spelgedrag komen of die stereotiep spelgedrag vertonen.

Op grond van bovengenoemde literatuur zijn zes stimuleringsgedragingen te onderscheiden. Deze zijn in te delen in minder en meer intensieve stimulering van het spelgedrag van kinderen:

Algemeen spelstimulerend gedrag:

- spelidee aanreiken voorafgaande aan het spel;
- aanwezig zijn en toekijken bij het spel (niet interveniëren);
- informatieve vragen stellen met het doel om kinderen te motiveren (zonder hulp of aanwijzingen te geven);
- waardering uitspreken voor hetgeen kinderen doen (zonder hulp of aanwijzingen te geven).

Directe sturing in spelstimulerend gedrag:

- extra sturing geven bij de uitvoering van spel; de leerkracht is gericht op correctie zonder een negatief oordeel uit te spreken over het spel van kinderen (bijvoorbeeld door 'waarom' vragen te stellen, die corrigerend zijn bedoeld of door mee te spelen of een demonstratie te geven);
- expliciete negatieve feedback; de leerkracht spreekt een negatief oordeel uit en probeert de uitvoering van het spel bij te sturen door extra hulp.

Hoe vaak en hoe intensief spelstimulering in de klassepraktijk wordt toegepast, is tot op heden niet onderzocht. In de huidige praktijk van groep één en twee in het basisonderwijs en wellicht ook in groep drie zullen spelstimuleringsactiviteiten wel voorkomen, maar naar verwachting niet zo intensief en systematisch als in spelstimuleringsexperimenten.

Ervaring van de leerkrachten

In het kleuteronderwijs zijn er soms grote verschillen in onderwijservaring tussen leerkrachten met groep 1 en 2. Het valt te verwachten dat ervaren leerkrachten de klasseorganisatie vlotter laten verlopen dan minder ervaren leerkrachten en daarom ook meer tijd beschikbaar hebben om tijdens de les leerlingen te stimuleren. Het aantal jaren ervaring in groep 1 en 2 is waarschijnlijk een relevante controlevariabele om rekening mee te houden bij het vaststellen van effecten van spelstimulering.

Klassemanagement

De mate waarin het spel door een kleuterleerkracht wordt gestimuleerd zal mede afhankelijk zijn van de klasseorganisatievaardigheden (het klassemanagement) zoals zij die in haar omgang met kleuters heeft eigen gemaakt. Managementvaardigheden liggen vooral op het terrein van het goed kunnen spreiden van aandacht, het op een juiste wijze en op het juiste moment ingrijpen bij ordeverstoringen of bij niet-actief spel of werkgedrag van kinderen en het tot een minimum beperken van lesonderbrekingen (zie Kounin, 1970; Doyle, 1986).

In dit onderzoek worden twee algemene kenmerken bestudeerd, namelijk in hoeverre de leerkracht gedurende de schooldag tijd nodig heeft om van de ene lesactiviteit (bijv. buitenspel) tot de andere lesactiviteit (bijv. speelwerkles) te komen en hoeveel tijd zij tijdens de speelwerkles besteedt aan klassemanagement (materialen of kinderen organiseren en ingrijpen bij storingen in de activiteiten van kinderen).

Sociaal-economische milieu-achtergronden van kinderen

Het komt nogal eens voor dat kinderen uit lagere sociaal-economische klassen minder geconcentreerd zijn en minder gericht op het uitvoeren van (zelfgestelde) opdrachten dan kinderen uit hogere milieus.

Het is daarom niet uitgesloten dat leerkrachten in klassen met kinderen uit milieus met een overwegend lagere sociaal-economische status (en vaak wat minder interesse voor schoolactiviteiten) meer tijd zullen besteden aan het management van de klas dan leerkrachten in klassen met kinderen uit milieus met een overwegend hogere sociaal-economische status (Brophy & Good, 1986). Daardoor zullen leerkrachten in deze klassen wellicht minder tijd en aandacht kunnen besteden aan het stimuleren van het spel van kinderen.

Onderzoeksvragen en verwachtingen

De volgende onderzoeksvragen zijn gesteld: *Onderzoeksvraag 1*: Welke verschillen zijn er in hoeveelheid tijd die kleuters besteden aan spelen enerzijds en aan werken en leren anderzijds en welke verschillen zijn er in de mate waarin kleuterleerkrachten met een KLOS- of PABO-opleiding het spelgedrag van kinderen stimuleren?

In de voorafgaande paragrafen is uiteengezet dat er reden is tot ongerustheid over de mate waarin kleuters gelegenheid wordt geboden om te spelen en over de mate waarin kleuterleerkrachten actief het spelgedrag van kleuters stimuleren. Ten aanzien van de eerste onderzoeksvraag zijn er de volgende verwachtingen. (1) Verwacht wordt dat in de huidige praktijk van het kleuteronderwijs stimulering van kinderen in spelsituaties wel zal voorkomen maar minder intensief zal zijn in vergelijking tot de stimulering van kinderen in werksituaties. Kleuterleerkrachten zullen doorgaans vanuit hun traditie met betrekking tot de vrije ontwikkeling van het kind niet sterk geneigd zijn om zich met het spel van kleuters te bemoeien (zie Wolbert, Schaap-Hummel & Span, 1986, over de opvattingen van kleuterleidsters). (2) Over verschillen in de mate van spelstimulering tussen leerkrachten met een KLOS- of een PABO/PA-opleiding zijn dan ook geen duidelijke verwachtingen geformuleerd.
Onderzoeksvraag 2: Hoe stimuleren leerkrachten in groep twee het spel van kinderen en in hoeverre beïnvloeden zij het spelgedrag van kinderen. Het spelgedrag is geoperationaliseerd als het niveau van betrokkenheid van een kleuter in een spelsituatie.

Deze vraag heeft betrekking op mogelijke effecten van spelstimulering op het spelgedrag van kinderen. Bij de beantwoording van deze vraag moet er rekening mee worden gehouden dat de algemene vaardigheid van leerkrachten in het organiseren van een klas (klassemanagement), de onderwijservaring van de leerkrachten en het sociaal milieu van de groep leerlingen condities vormen die de effecten van spelstimulering mede bepalen.

Er zijn drie verwachtingen naar aanleiding van de tweede onderzoeksvraag.

1. Leerkrachten in klassen met kinderen uit lagere sociale milieus zullen waarschijnlijk meer tijd besteden aan klassemanagement en daardoor minder tijd besteden aan spelstimulering dan leerkrachten met klassen met een hogere milieu-samenstelling.
2. Leerkrachten die veel tijd aan spelstimulering besteden en leerkrachten die weinig tijd nodig hebben voor klassemanagement zullen bij de kinderen een grote spelbetrokkenheid uitlokken.
3. Leerkrachten met ervaring in het kleuteronderwijs zullen minder tijd nodig hebben voor de klasseorganisatie en minder tijd nodig hebben voor de overgang van de ene naar de andere lesactiviteit.

Variabelen en instrumenten

In Tabel 1 is een overzicht gegeven van de variabelen die in het onderzoek zijn opgenomen en van de instrumenten waarmee de variabelen zijn gemeten. Er zijn drie instrumenten gehanteerd: (1) een leerkrachtenvragenlijst om achtergrondvariabelen van de leerkracht en de klas vast te stellen, (2) een lesobservatie-instrument om het gedrag van leerkrachten en leerlingen tijdens twee speelwerklessen vast te leggen en (3) een dagrooster waarin gedurende een gehele schooldag de activiteiten van de kinderen per tijdseenheid zijn genoteerd door de observator. De observatieprocedure is als volgt verlopen.

De activiteiten in een kleuterklas zijn gedurende één schooldag in kaart gebracht. Er zijn observaties gedaan van twee speelwerklessen tijdens de schooldag. Geobserveerd zijn de spelstimuleringsactiviteiten en klassemanagementactiviteiten van de leerkrachten en het niveau van betrokkenheid van groepjes kinderen in verschillende spel- of werksituaties geobserveerd.

Ten einde de *betrokkenheid van kinderen bij hun spel* en hun werk in de klas vast te kunnen stellen is het gedrag van de kinderen in de diverse hoeken en bij de verschillende tafels tijdens speel/werklessen geobserveerd. Het gedrag van elk van de kinderen in een groepje bij een bepaalde tafel of hoek werd gecodeerd in één van vier categorieën. Deze vier categorieën zijn geoperationaliseerd in termen van vier niveaus van gerichte en aangehouden aandacht voor de momentane activiteit (zie voor een be-

schrijving p. 339). Op een volgend meetmoment werd de volgende hoek of tafel geobserveerd. Deze cycli werden herhaald tot het einde van de speel/werkles.

Aangezien tijdens deze observaties niet alleen betrokkenheid van kinderen werd gecategoriseerd maar tevens werd vastgelegd of ze betrokken waren bij spel, werk of leeractiviteiten, kon worden vastgesteld:

- gelegenheid tot spel: het percentage van de lestijd dat het gemiddelde kind doorbrengt in spelsituaties tijdens de speel/werkles;
- gelegenheid tot werken of leren: het percentage van de lestijd dat het gemiddelde kind bij werkhoeken/tafels doorbrengt;
- spelbetrokkenheid: de gemiddelde betrokkenheid van kinderen bij hun activiteiten in spelsituaties;
- betrokkenheid bij werk- of leersituaties: de gemiddelde betrokkenheid van kinderen bij werk- of leersituaties.

Het *gedrag van de leerkracht* werd eveneens herhaaldelijk vastgelegd tijdens de speel/werkles. Er zijn zes categorieën onderscheiden teneinde de volgende percentages te kunnen berekenen:

- afwezig/nietsdoen: het percentage van de lestijd dat de leerkracht afwezig is of geen aandacht voor de kinderen heeft, rond loopt zonder gericht te zijn op de leerlingen of de klassesituatie;
- managementsgedrag: het percentage van de lestijd dat de leerkracht besteedt aan klassemanagement tijdens de speel/werklessen;
- algemene spelstimulering: het percentage van de tijd dat de leerkracht aan algemene stimulerende acties besteedt;
- directe sturing: het percentage van de tijd voor directe sturing van spelende kinderen in de speel/werklessen;
- algemene stimulering in werk- of leersituaties: het percentage van de tijd dat de leerkracht aan algemene stimulering van kinderen in de werkhoeken of in leersituaties besteedt;
- directe sturing werk- of leersituaties: het percentage van de tijd voor directe sturing van kinderen in de werkhoeken of leersituaties.

Observatieprocedure

De observatie van leerling- en leerkrachtgedrag werd georganiseerd in de vorm van een 'rotating time sample' waarbij in elke cyclus van twee minuten éénmaal een groepje leerlingen in een hoek of bij een tafel en éénmaal de leerkracht werd geobserveerd. Omdat de tijdsbasis van de afzonderlijke observatiemomenten exact vastgelegd is, is het niet mogelijk dat er interactie is opgetreden tussen het gecategoriseerde gedrag en het aantal malen dat dit gedrag gecategoriseerd is. Deze voorzorgsmaatregel is van belang wanneer bepaald gedrag (bijv. nietsdoen) gemakkelijker herkend wordt en derhalve sneller gecodeerd kan worden dan complexer gedrag (bijv. sturing).

Ten einde tijdens de training van observatoren te kunnen schatten of de observatoren voldoende betrouwbaar waren wat betreft de beoordeling van leerling- en leerkrachtgedrag, hebben de vijf observatoren één op videoband vastgelegde speel/werkles overeenkomstig de voorgeschreven procedures beoordeeld. Deze proef resulteerde voor elke observator in 14 oordelen over het leerlinggedrag op 14 verschillende momenten tijdens de speel/werkles en in 15 oordelen op 15 verschillende momenten over het leerkrachtgedrag.

De oordelen over het gedrag van leerlingen werden overeenkomstig de te analyseren variabelen met betrekking tot de betrokkenheid van leerlingen getransformeerd naar de gemiddelde betrokkenheid van de leerlingen die op een bepaald moment geobserveerd werden. Aangezien door deze procedure intervaldata gegenereerd worden, is de betrouwbaarheid van de observatoren wat betreft de beoordeling van leerlinggedrag geschat aan de hand van de variantie-analytische benadering (Winer, 1961). Deze benadering houdt in dat de verschillen tussen de beoordeelde objecten (is gemiddelde betrokkenheid op verschillende momenten tijdens de les) worden afgezet tegen de verschillen tussen de beoordelaars per beoordeeld object. Dit betekent dat de betrouwbaarheid van de oordelen hoger is naarmate de observatoren het meer met elkaar eens zijn en naarmate de beoordeelde objecten sterker van elkaar verschillen.

Berekend werd de betrouwbaarheid van de gemiddelde beoordelaar zonder correctie voor ankerpunten omdat de onderzoeksprocedures

impliceren dat elke klas door een enkele observator beoordeeld wordt. Deze betrouwbaarheid bleek r_1=.73, hetgeen voldoende hoog is om op betrouwbare wijze de in het onderzoek geoperationaliseerde variabelen te meten.

De categorieën waarin het leerkrachtgedrag gecodeerd werd, zijn ten behoeve van de schatting van de betrouwbaarheid van de beoordeling van het leerkrachtgedrag teruggebracht tot vier categorieën die van belang zijn voor de constructie van de variabelen met betrekking tot het leerkrachtgedrag (sturing, algemene stimulering, management en rondlopen/niets doen). Gezien het nominale meetniveau van deze data is Cohens kappa gebruikt voor de schatting van de betrouwbaarheid. De kappa voor de gemiddelde observator was .75, een voldoende hoge interbeoordelaarsbetrouwbaarheid (zie Popping, 1983).

Steekproefontwerp

Gezien het beperkte budget voor het onderhavige onderzoek is er voor gekozen om de steekproef te beperken tot Noord-Nederland. Een landelijke steekproef zou door het aantal reisuren en de reis- en verblijfkosten de grootte van de steekproef beperken tot ± 40. De resulterende steekproefruis werd schadelijker geacht

Tabel 1

Variabelen en instrumenten

		leerkracht-vragenlijst	lesobservatie-instrument	observatie dagrooster
Achtergrondvariabelen				
a.	Sociaal-economische status van leerlingen in een klas	x		
b.	Vooropleiding van de leerkracht van groep 2: KLOS of PA/PABO	x		
c.	Aantal jaren onderwijservaring in groep 1 of 2	x		
d.	Weekrooster: tijd tijdens een schoolweek besteed door kinderen aan speel-, werk- en leeractiviteiten in groep 2	x		
Klassemanagement en overig algemeen gedrag				
e.	Discontinuïteit in onderwijstijd: tijd die kinderen tussen de lesactiviteiten besteden aan organisatorische activiteiten (verkleden, aan het werk gaan etc.)			x
f.	Klasse-organisatie tijdens de speelwerkles (tijd door de leerkracht besteed aan organisatie en orde handhaven tijdens de les)		x	
g.	afwezig/niets doen (tijd die niet aan de leerlingen of de klassesituaties wordt besteed)		x	
Spelinterventies				
h.	Algemeen stimulerend lesgedrag: leerkrachtgedrag gericht op het stimuleren van de betrokkenheid van kinderen (geen directe sturing)		x	
i.	Directe sturing: feedback en demonstratie door de leerkracht gericht op verbetering van uitvoeringsgedrag van kinderen		x	
Gelegenheid				
j.	Gelegenheid tot spelen of werken: het percentage lestijd dat het gemiddelde kind doorbrengt in spel, werk of leersituaties tijdens de speelwerkles		x	
Spelgedrag				
k.	Niveau van betrokkenheid van kinderen in spelsituaties		x	

voor het trekken van valide conclusies uit de te verzamelen data dan een tweemaal grotere steekproef die vermoedelijk slechts licht vertekend is door regionale verschillen. Vooral voor relaties tussen kenmerken van leerkrachten en leerlingen geldt dat regionale verschillen weinig plausibel zijn. Voor verdelingskenmerken van enkele variabelen die van weinig inhoudelijk belang zijn, maar als controlevariabele in het onderzoek opgenomen zijn zoals de gemiddelde sociaal-economische status van een klas, mag aangenomen worden dat de verdeling voor met name West-Nederland afwijkt van die van Noord-Nederland.

Een tweede beperking van de doelpopulatie is dat alleen leerkrachten in de doelpopulatie worden opgenomen die uitsluitend een volledige groep twee van de basisschool leiden. Het argument voor deze beslissing is vergelijkbaar met de argumentatie voor de hiervoor beschreven geografische beperking. Combinatiegroepen bestaan uit leerlingen die veel sterker dan leerlingen in groep twee variëren in leeftijd. Om de resulterende ruis te controleren zou de variabele leeftijd in het onderzoek betrokken moeten worden hetgeen de complexiteit en daarmee de kosten van de observaties van leerlinggedrag ontoelaatbaar zou verhogen. Daarnaast geldt het argument dat groep twee inhoudelijk van groter belang is dan groep één met het oog op de doelstelling van het onderzoek met betrekking tot de vraag naar de verschraling van spelaanbod en -stimulering in de basisschool. In dat kader is de overgang van groep twee naar groep drie met het oog op de (vroegere?) strikte scheiding tussen kleuter- en lager onderwijs interessanter dan van andere combinaties van groepen.

De eenheden werden getrokken door een eenvoudige aselecte trekking uit die scholen voor basisonderwijs in Noord-Nederland die wat betreft het aantal leerlingen in de kleuterklassen vermoedelijk niet noodzakelijk combinatieklassen hoeven samen te stellen. Daartoe is op basis van gegevens van het Ministerie van Onderwijs en Wetenschappen een selectie samengesteld van scholen die in het voorafgaande schooljaar tenminste 25 kinderen in groep één hadden. Van de 230 scholen die in eerste instantie aan de genoemde criteria voldoen, bleek echter ongeveer de helft na telefonische bevraging toch niet bruikbaar omdat een klas met uitsluitend groep twee-leerlingen ontbrak omdat bijvoorbeeld parallelle combinatieklassen geprefereerd werden of omdat de school te weinig leerlingen in groep twee bevatte. Van de resterende scholen werkten uiteindelijk 62 scholen mee. Doordat op enkele scholen sprake was van parallelle groepen twee konden uiteindelijk 65 observatie-protocollen gecreëerd worden.

Er werd niet disproportioneel gestratificeerd naar kenmerken van scholen of leerkrachten. Aangezien de doelstelling deels een beschrijving inhoudt wat betreft de status quo ten aanzien van spelstimulering en de gelegenheid tot spel werd het behoud van de natuurlijke variantie in de verzamelde data geprefereerd boven de precisie in de schattingen voor specifieke subgroepen die gezien de omvang van de totale steekproef toch al beperkt is.

Resultaten

Tijd om te spelen en tijd om te werken in groep 2

Achtergrondkenmerken van leerlingen en leerkrachten in groep 2 zijn dat van de 65 leerkrachten 53 leerkrachten een KLOS-opleiding hebben gevolgd en 12 leerkrachten een PA of PABO-opleiding hebben. Het gemiddeld aantal jaren ervaring in groep 1 en 2 van de leerkrachten is 10 jaar (*sd* = 7). Het gemiddeld aantal kinderen per klas is 26 (*sd* = 4) en de sociaal-economische status van de kinderen in de klassen is als volgt te typeren: 8% hoge status, 83% gemiddelde of gemengde status en 9% lage sociale status.

Gelegenheid om te spelen

In groep 2 is globaal genomen de verdeling van de totale schooltijd als volgt: de kinderen besteden 11% van de totale schooltijd aan organisatie (verkleden, binnenkomen, taakverdelen etc), 28% van de tijd aan kringactiviteiten (gesprek, pauzehapje, kringspel, zang, voorlezen etc.), 9% van de tijd aan voorbereidend lezen, schrijven of rekenen met hele groep of in subgroepjes, 29% van de tijd aan spel in gymnastiekzaal of buiten, 10% aan spelsituaties in de klas en 13% aan werksituaties in de klas. Kleuters besteden dus doorgaans circa 39% van de *totale schooltijd* in spelsituaties. Gedurende de

Tabel 2

Percentage van de tijd in de speel/werkles dat kinderen gemiddeld in spel-, werk- of leersituaties aanwezig zijn

Spelsituaties		Werksituaties		Leersituaties	
– watertafel	2%	– kleien	2%	– lees/schrijfoefeningen	1%
– zandtafel	1%	– ontwik.mat.	12%	– rekenactiviteiten	1%
– constructietafel	9%	– tekenen, schilderen	20%	– werkbladen	4%
– bouwhoek	10%	– knippen, plakken, prikken, vlechten, e.d.	20%		
– timmerhoek	0,5%	– overige activiteiten	4%		
– poppenhoek	9%				
– weeghoek	0%				
– winkeltje	0,5%				
– poppenkast	0,5%				
– leeshoek/ luisterhoek	2,5%				
	35%		58%		6%

twee geobserveerde speelwerklessen in een schooldag is nagegaan hoeveel tijd kinderen doorbrengen in de verschillende spelsituaties. Tabel 2 geeft een overzicht.

Gemiddeld genomen brengen kinderen 35% van de tijd van de speelwerklessen door in spelsituaties, 58% in werksituaties en 6% in leersituaties (meestal gaat het bij dit laatste om het maken van werkbladen met voorbereidende lees- en schrijfoefeningen). Leerkrachten met een KLOS-opleiding verschillen niet van leerkrachten met een PA- of PABO-opleiding wat betreft het percentage lestijd dat kinderen in spelsituaties tijdens de speel/werkles doorbrengen ($F(1,60) = 2.4; p = .13$) of in het aantal verschillende spelsituaties dat een klas tijdens een schoolweek krijgt aangeboden ($F(1,46) = 1.07; p = .30$). Deze twee groepen van leerkrachten verschillen overigens ook niet in het percentage kinderen dat tijdens de speel/werkles in werksituaties is ($F(1,60) = 0.02; p = .88$) of in leersituaties is ($F(1,60) = 0.005; p = .94$). In de spelsituaties zijn de kinderen 22%

Tabel 3

Leerkrachtgedrag tijdens speel/werklessen (in percentages van de lestijd van de geobserveerde speel/werklessen)

		Totaal	KLOS	PA/PABO
Algemene lesgedragingen				
1.	afwezig of geen aandacht	14,0%	14,5%	13,5%
2.	loopt rond	9,5%	10,5%	9,0%
3.	klassemanagement	30,5%	30,5%	29,5%
4.	**Algemeen spelstimulerend gedrag**	6,5%	6,5%	5,0%
4.1	kijkt toe bij het spel	(3,0)	(3,5)	(1,5)
4.2	geeft spel idee vooraf	(2,5)	(2,0)	(2,5)
4.3	stelt informatieve vragen	(0,5)	(0,5)	(0,5)
4.4	spreekt positieve waardering uit	(0,5)	(0,5)	(0,5)
5.	**Directe sturing bij spelsituaties**	2,5%	2,0%	1,5%
5.1	geeft extra hulp	(2,0)	(1,5)	(0)
5.2	geeft negatieve feedback (beoordeling) en extra hulp	(0,5)	(0,5)	(1,5)
6.	**Stimulering van kinderen in werksituaties**	34,0%	31,0%	40,0%
6.1	algemeen stimulerend gedrag	(17,5)	(15,5)	(23,5)
6.2	directe sturing	(16,5)	(15,5)	(16,5)
7.	**Stimulering van kinderen in leersituaties**	3,5%	4,5%	0,5
7.1	algemeen stimulerend gedrag	(1,5)	(2,0)	(0)
7.2	directe sturing	(2,0)	(2,5)	(0,5)

Noot: de percentages tussen haakjes vormen een uitsplitsing van het er bovengenoemde percentage

van de lestijd bezig met constructiespel (watertafel, zandtafel, constructietafel, bouwhoek en timmerhoek), 10% met symbolisch spel (poppenhoek, poppenkast, winkeltje) en 2% met regelspel (kijk- en luisterspelletjes). Senso-motorisch spel vindt meestal plaats in het speellokaal of op de speelplaats buiten de school. De tijd in werksituaties tijdens de speelwerklessen wordt door de kinderen meestal besteed aan tekenen en schilderen, knippen en plakken en aan het werken met ontwikkelingsmateriaal (puzzels en lotto's). Er wordt doorgaans vrijwel geen tijd besteed aan het leren lezen, schrijven of rekenen.

Spelinterventies en de betrokkenheid van kinderen

Leerkrachten besteden relatief weinig tijd aan het stimuleren of sturen van kinderen in spelsituaties.

Tabel 3 laat zien dat gemiddeld genomen (kolom: Totaal) de leerkrachten 9% van de lestijd aandacht hebben voor spelsituaties (algemeen spelstimulerend 6,5% en directe spelsturing 2,5%). De verdeling tussen de leerkrachten is als volgt: 53% van de leerkrachten is tot 5% van de lestijd gericht op spelsituaties, 27% van de leerkrachten is 6% tot 15% van de lestijd gericht op spelsituaties en 20% van de leerkrachten is 16% tot 54% van de tijd gericht op spelsituaties. Tabel 3 laat verder zien dat leerkrachten 9% van de lestijd aandacht hebben voor spelsituaties en 34% van de tijd voor werksituaties en 3,5% van de tijd voor leersituaties. Met name in spelsituaties wordt naar verhouding weinig directe sturing gegeven door de leerkracht. Deze resultaten ondersteunen de assumptie in dit onderzoek dat kinderen in spelsituaties vrij worden gelaten in wat zij willen doen en hoe zij hun spel uitvoeren.

De verschillen in spelstimulering tussen leerkrachten met een KLOS-opleiding versus PA- of PABO-opleiding zijn statistisch niet significant. KLOS-opgeleiden geven dus niet méér algemene spelstimulering of directe sturing in spelsituaties dan leerkrachten met een PA/PABO-vooropleiding. (De statistische toetswaarden van de factoren 'percentage algemene spelstimulering tijdens de les' en 'percentage directe spelsturing tijdens de les' zijn respectievelijk F (1,58) = 0.10; p = .75 en F (1,58) = 0.002; p = .87.) De verschillen in managementsgedrag en in stimulering van werksituaties en in stimulering van leersituaties tussen KLOS- en PA/PABO-leerkrachten zijn eveneens niet statistisch significant.

Een opvallend gegeven is dat leerkrachten gemiddeld genomen slechts 9% van de lestijd gericht zijn op spelsituaties terwijl kinderen doorgaans 35% van de tijd in deze situaties doorbrengen. In deze speelwerklessen besteden leerkrachten 34% van hun tijd aan werksituaties waar kinderen 58% van de tijd doorbrengen. Leerkrachten zullen naar verhouding minder tijd besteden aan kinderen in spelsituaties dan aan kinderen in werksituaties.

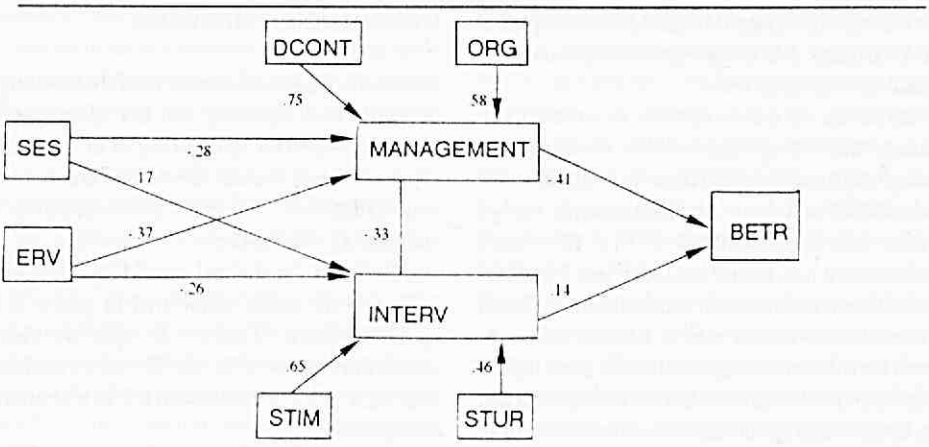
De betrokkenheid van kinderen is gemeten door observatie van het niveau van aangehouden en gerichte aandacht van kinderen in spelsituaties.

Tabel 4 laat zien dat over het geheel genomen tijdens de geobserveerde lessen 73% van

Tabel 4

Percentage leerlingen en hun niveau van betrokkenheid bij spelsituaties

Niveau 1:	**niet spelen**, het kind kijkt vrijwel de gehele geobserveerde tijd rond, is regelmatig van de plaats af, of praat over andere dingen dan het spel	13%
Niveau 2:	**oppervlakkige spelactiviteiten**, het kind kijkt voor de helft van de geobserveerde tijd rond of manipuleert materiaal zonder er iets doelgericht mee te doen	14%
Niveau 3:	**tamelijk intensieve activiteit**, het kind is het grootste gedeelte van de geobserveerde tijd bezig maar laat zich incidenteel afleiden (begint met anderen te praten, breekt spel af en begint opnieuw etc.)	23%
Niveau 4:	**zeer intensieve activiteit**, het kind laat zich zelden of niet afleiden, blijft actief met het spel bezig; de ogen zijn steeds op het materiaal of de medespelers gericht	50%



Figuur 1. Model van relaties tussen spelinterventies, klassemanagement en betrokkenheid van kinderen in spelsituaties

de kinderen tamelijk intensief tot intensief betrokken (niveau 3 en 4) is in de spelsituaties en 27% van de kinderen nauwelijks of niet bezig is (niveau 1 en 2) in de spelsituaties. Het gemiddelde niveau van betrokkenheid van de kinderen is berekend door het gemiddelde niveau van betrokkenheid bij spelsituaties per klas te middelen over alle klassen. Dit gemiddelde bedraagt 3.1 (*sd* 0.4). De minimum betrokkenheidsscore van een klas is 2.2 en de maximumscore van een klas 3.8. Er zijn dus grote verschillen tussen klassen in het niveau van betrokkenheid van de kinderen. Dit maakt het interessant om te kijken of verschillen in betrokkenheidsniveau tussen klassen gerelateerd kunnen worden aan verschillen in leerkrachtengedrag.

De relatie tussen spelstimulering en het niveau van betrokkenheid van kinderen in spelsituaties

Er is een correlationele analyse uitgevoerd. De drie algemene verwachtingen over de relatie tussen spelinterventies en klassemanagement enerzijds en het niveau van betrokkenheid van kinderen anderzijds zijn gespecificeerd in een beschrijvend model dat is beproefd met behulp van LISREL (Jöreskog & Sörbom, 1985).

De basisassumptie van het model is dat spelinterventievariabelen en klassemanagementvariabelen het niveau van betrokkenheid van kinderen in spelsituaties voorspellen. In Figuur 1 wordt het model gepresenteerd zoals ten slotte is beproefd. In dit model is de factor opleiding (KLOS versus PA/PABO) van de leerkracht niet opgenomen omdat deze niet samenhing met de spelinterventies en het klassemanagement of met de spelbetrokkenheid van kinderen. Deze samenhangen werden van te voren ook niet verondersteld.

In het model worden de relaties aangeduid met pijlen waarbij partiële correlatiecoëfficiënten zijn vermeld. De volgende onafhankelijke variabelen zijn opgenomen: Sociaal-economische status van de klas (SES: lagere status=1, gemiddelde en gemengde status=2 en hoge status=3) en aantal jaren ervaring (ERV) van de leerkracht in groep 1 en 2. Er zijn twee procesfactoren: spelinterventies (INTERV) en management van de klas. De eerste factor is samengesteld uit twee indicatoren: het percentage lestijd door de leerkracht besteed aan stimuleren van het spel (STIM) en het percentage lestijd besteed aan sturen van het spel (STUR). De tweede factor is samengesteld uit de indicatoren: discontinuïteit in lestijd van leerlingen (percentage van de schooltijd die kinderen tussen de lesactiviteiten verliezen: DCONT) en percentage van de lestijd dat leerkrachten aan klasseorganisatie besteden (ORG). De onafhankelijke variabele is het niveau van betrokkenheid van kinderen in spelsituaties (BETR). Het model is beproefd en de resultaten wijzen uit dat de oorspronkelijke correlatiematrix en de door het model geschatte matrices goed met elkaar in overeenstemming zijn (Chi-kwadraat met 9 vrijheidsgraden= 5.0; p=.83 en de goodness of fit index=.93. De hoogste waarde van

de residuen bedraagt 0.07, hetgeen acceptabel is). In Bijlage 1 is de oorspronkelijke correlatiematrix weergegeven.

De *eerste verwachting* over de invloed van management en spelinterventies wordt slechts gedeeltelijk ondersteund door de resultaten. De padcoëfficiënt tussen managementtijd en het niveau van betrokkenheid is -.41. Hoe meer leerkrachten hun lestijd besteden aan klasseorganisatie en hoe meer tijd de kinderen verliezen tussen lesactiviteiten des te minder zijn kinderen betrokken in de spelsituaties. Daar tegenover staat dat de tijd die leerkrachten gebruiken om kinderen in spelsituaties te stimuleren of te sturen geen significante relatie heeft met de spelbetrokkenheid van kinderen (padcoëfficiënt is .14). De tijd benodigd voor klassemanagement is kennelijk dermate bepalend voor spelinterventies door de leerkracht (padcoëfficiënt is -.33) dat de oorspronkelijke positieve en significante correlatie tussen spelinterventies en spelbetrokkenheid er deels door verklaard wordt.

De *tweede verwachting* over de invloed van de sociaal-economische status op de managementtijd en via deze factor op de betrokkenheid van kinderen wordt ondersteund door de resultaten. Tijdens de analyse bleek het nodig om ook een pad van de sociaal-economische status naar spelinterventies te trekken. Een hogere sociaal-economische status van kinderen in een klas kan zowel via minder tijd voor klassemanagement als via meer spelinterventies van de leerkracht van invloed zijn op de betrokkenheid van kinderen.

De *derde verwachting* over de invloed van ervaring van leerkrachten met groep 1 en 2 op de benodigde tijd voor klassemanagement wordt eveneens door de resultaten ondersteund. Het bleek bovendien nodig om een pad te trekken van ervaring naar spelinterventies. Leerkrachten met veel ervaring in groep 1 en 2 blijken niet alleen beduidend minder tijd te besteden aan klassemanagement maar ook aan spelinterventies (en dan met name aan directe sturing van kinderen in spelsituaties). Opvallend is dat het hier gaat om leerkrachten met een kleuteropleiding die een ruim aantal jaren (meer dan 10 jaar) ervaring hebben. Wellicht dat de filosofie uit de zeventiger jaren over de vrije en natuurlijke ontwikkeling van kinderen deze leerkrachten weerhoudt van bemoeienis met het spel van kinderen.

Conclusie en discussie

De *eerste onderzoeksvraag* heeft betrekking op de aard en de omvang van het spelaanbod in kleutergroepen en op de vraag of er verschillen zijn in de mate waarin kleuterleerkrachten met een KLOS- of een PABO/PA-opleiding het spelgedrag van kinderen stimuleren.

Kinderen besteden gemiddeld genomen 39% van de totale schooltijd in groep 2 aan spelactiviteiten. Tijdens de speelwerklessen wordt doorgaans 35% van de tijd in spelsituaties, 58% in werksituaties en 6% in leersituaties doorgebracht.

Kinderen in spelsituaties brengen de meeste tijd door in de bouwhoek, de poppenhoek en aan de constructietafel. Andere spelsituaties zoals de zand- en watertafel, het winkeltje, de weeghoek etc. die in handboeken voor het kleuteronderwijs worden beschreven (zie Janssen-Vos, Dikken & Verkley, 1990) worden aanzienlijk minder gebruikt of zijn niet aanwezig in de school.

Het spelaanbod dat doorgaans in groep 2 wordt gegeven is, gezien de geringe variatie in mogelijkheden (zie Tabel 2) en de beperkte tijd die kinderen in spelsituaties doorbrengen voor verbetering vatbaar. Wellicht dat bij een verbetering van het spelaanbod ook een verhoging van de spelbetrokkenheid van kinderen is te bewerkstelligen.

De onderzoeksresultaten laten zien dat kleuterleerkrachten relatief weinig stimulering geven aan kinderen in spelsituaties en naar verhouding meer stimulering geven aan kinderen in werk- of leersituaties.

Er zijn geen statistisch significante verschillen geconstateerd tussen leerkrachten met een PABO/PA-opleiding of met een KLOS-opleiding in tijd die kinderen in spelsituaties doorbrengen of in de variatie in spelsituaties in de school. Leerkrachten met een KLOS-opleiding verschillen evenmin van de PABO/PA-opgeleiden in de aard en hoeveelheid spelstimulering voor kleuters (zie Tabel 3).

Opvallend is het vrij hoge percentage van de lestijd dat leerkrachten afwezig zijn of geen aandacht hebben voor de kinderen (ca. 14% van de tijd).

De *tweede onderzoeksvraag* heeft betrekking op de relatie tussen het lesgeefgedrag en het betrokkenheidsniveau van kinderen in spel-

situaties. In het lesgeefgedrag zijn twee aspecten onderscheiden: klassemanagement en spelstimulering. Daarnaast is bezien of de factor onderwijservaring een rol speelt in het lesgeefgedrag.

Ons onderzoek leidt tot de conclusie dat de managementfunctie van het onderwijzen, in het bijzonder voor het bevorderen van de betrokkenheid van jonge kinderen in spelsituaties van groot belang is. Minder ervaren leerkrachten besteden tijdens de speelwerkles meer tijd aan het management dan ervaren leerkrachten. Daardoor hebben minder ervaren leerkrachten waarschijnlijk kinderen met een lager niveau van betrokkenheid in hun klas. Minder ervaren leerkrachten in groep 2 kunnen wellicht worden getraind in het efficiënt leren organiseren van de speel- en werk- en leeractiviteiten van kinderen. Efficiënt klassemanagement impliceert niet alleen het leren inrichten van een bruikbare en gevarieerde speelwerkomgeving, maar ook het leren toepassen van een systeem van gedragsregels dat het niveau van betrokkenheid van kinderen bij een activiteit hoog houdt (zie Kounin, 1970). Ervaren kleuterleerkrachten bezitten deze vaardigheden in hogere mate, maar deze groep kleuterleerkrachten besteedt weinig aandacht aan spelstimulering tijdens de speelwerklessen. Verwacht mag worden dat wanneer ervaren kleuterleerkrachten meer spelstimulering aan kinderen gaan geven de betrokkenheid van kinderen bij hun spelsituaties zal worden vergroot. De ervaren kleuterleerkrachten dienen op de hoogte te worden gebracht van de mogelijkheden van spelinterventies om de betrokkenheid van kinderen te vergroten. Programma's voor spelen in en rond de school, met spelactiviteiten voor verschillende groepen kinderen en die van dag tot dag kunnen worden toegepast, zullen ervaren kleuterleerkrachten een basis verschaffen om de kleuters meer te stimuleren in hun spel. Bij leerkrachten is behoefte aan dergelijke programma's (zie Harskamp, Pijl & Snippe, 1991). Er zijn in het verleden en recentelijk enkele aanzetten gegeven voor zulke programma's (zie *Speciale Spelbegeleiding*, Bosch & Bruins, 1984; *Spel en Spelen*, Janssen-Vos, Dikken & Verkley, 1990). In toekomstige onderzoeksprojecten ten behoeve van het kleuteronderwijs kan de ontwikkeling en implementatie van spelprogramma's worden onderzocht. Vooral spelprogramma's die gericht zijn op de ontwikkeling van algemene basisvaardigheden (zoals betrokkenheid bij het uitvoeren van taken) en van specifiekere vaardigheden (zoals het zich ontwikkelend lezen, schrijven en rekenen) lijken in de hedendaagse basisschool bijzonder welkom (zie Harskamp & Willemsen, 1991; Verhoeven & Mommers, 1992).

Daarnaast is onderzoek naar de ontwikkeling en implementatie van klassemanagementcursussen voor de onderbouw van de basisschool van groot belang voor het verhogen van de betrokkenheid van kinderen bij beginnende en onervaren leerkrachten. Het Christelijk Pedagogisch Studiecentrum zal cursusjaar 1993/1994 op een aantal PABO's met een dergelijke cursus starten.

Literatuur

Berlyne, D. (1960). *Conflict, Arousal and curiosity.* New-York: MacGraw-Hill.

Bosch, N., & Bruins, A. (1984). *Speciale spelbegeleiding werkboek.* Tilburg: Zwijsen.

Brophy, J. E., & Good (1986). Teacher behavior and student achievement. In M. C. Wittrock, (Ed.), *Handbook of research on teaching* (pp. 328-375). New York: MacMillan.

Bühler, Ch. (1928). *Kindheit und Jugend.* Leipzig.

Christie, J. F., & Johnson, E. P. (1983). The role of play in social intellectual development. *Review of Educational Research*, Spring, vol. 5, 3, 93-115.

Doyle, W. (1986). Classroom organisation and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 392-431). New York: MacMillan.

Duim, G. J. van der, & Blok, H. (1990). Knelpunten in de opleiding tot kleuterleerkracht. *Didaktief*, 20(8), 20-22.

Erikson, E. (1974). *Childhood and society.* New York: Norton.

Faber, S. E. K. (1990). De effecten van spelbegeleiding in het reguliere en het speciale onderwijs. *Tijdschrift voor Orthopedagogiek, 29*, 84-108.

Harskamp, E., & Willemsen, T. (1991). Programma's en speelleermaterialen voor voorbereidend rekenen in de basisschool. *Pedagogische Studiën, 68*, 404-415.

Harskamp, E., Pijl, S.J., & Snippe, J. (1991). *Spel en spelstimulering in het primair onderwijs*. Groningen: RION.

HBO-raad (1989). *Actieplan lerarenopleiding basisonderwijs*. Hogeschool Bericht, nr. 6, 8. 151-166.

Hellendoorn, J., & Zanten, B.L. van (1989). Speltraining bij basisschoolleerlingen, eerste resultaten. *Nederlands Tijdschrift voor Onderwijs en Opvoeding, 5*(2), 76-86.

Janssen-Vos, F., Dikken, N., & Verkley, H. (1990). *Spel en spelen, kleuters in de basisschool*. Assen: Van Gorcum.

Janssen-Vos, F. (1991). Van kleuteronderwijs naar basisonderwijs, wat zich in tien jaar tijd afspeelde. *De Wereld van het Jonge Kind, 8*(6) februari, 140-147.

Jöreskog, K.G. & Sörbom, D. (1985). *Analysis of Lineair Structural Relationship by the method of maximum likelihood*. Users' guide LISREL VI. Upsala: University

Kooy, R. van der, & Rost, D.M. (1985). Spelonderzoek en spelstimulering. *Pedagogisch Tijdschrift*, 138-148.

Kopp, C., & Vaughn, B. (1982). Sustained attention during explorary manipulation as a predictor of cognitive competence in preterm infants. *Child Developtment*, 52, 174-182.

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

Kuyk, J.J. van (1987). Basisvorming in het kleuteronderwijs. In J.J. van Kuyk (red.), *Basisvorming in de basisschool* (pp. 219-240). Tilburg: Zwijsen.

Laevers, F. (1991). *The innovative project Experiental Education*. Leuven: Katholieke Universiteit.

Levy, A.I.C. (1984). The language of play. The role of play in language development, a review of literature. *Early Child Development and Care, 17*, 49-62.

Piaget, J. (1962). *Play, dreams and imitation*. New York: Norton.

Popping, R. (1983). *Overeenstemmingsmaten van nominale data*. Academisch proefschrift Rijksuniversiteit Groningen.

Rubin, K.H., Fein, G., & Vandenberg, B. (1983). Play. In P. Mussen (Ed.), *Handbook of child psychology, vol. 4* (pp. 693-774). New York: John Wiley.

Ruff, H.A., & Lawson, K.R. (1990). Development of sustained and focused attention in young children during free play. *Developmental Psychology, 26*(1), 85-93.

Singer, J.L. (1985). Verbeeldend spel: implicaties voor preventie en kinderpsychotherapie. In J. Hellendoorn (red.), *Therapie, kind en spel* (pp. 89-107). Deventer: Van Loghum Slaterus.

Ven, J. van de, Vrieze, G., & Claessen, J. (1988). *Grensverkenners in de basisschool. Een onderzoek naar volledige inzetbaarheid van voormalige LO- en KO-leerkrachten in het basisonderwijs*. Nijmegen: ITS.

Verhoeven, L., & Mommers, C. (1992). Ontwikkelings- en programmagericht leesonderwijs. In L. Verhoeven (red.). *Handboek lees- en schrijfdidactiek* (pp. 101-118). Amsterdam/Lisse: Swets & Zeitlinger.

Winer, B.J. (1961). *Statistical principles in experimental design* (2nd. ed.). London: McGraw-Hill.

Wolbert, R., Schaap-Hummel, J., & Span, P. (1986). *Individualisering en differentiatie in de basisschool*. 's-Gravenhage: SVO.

Manuscript aanvaard 7-6-1993

Auteur

E.G. Harskamp is als coördinator onderwijsonderzoek verbonden aan het RION, Instituut voor Onderwijsonderzoek te Groningen.

Adres: RION, Instituut voor Onderwijsonderzoek, Westerhaven 15, 9718 AW Groningen.

Abstract

Play at school

E.G. Harskamp. Pedagogische Studien, 1993, 70, 337-351.

Six years after the implementation of the Primary Education Act in the Netherlands, a controversy has arisen concerning the quality of the educational programme for four- to six-year-olds. According to educational experts, the new programme is in danger of losing play intervention as an important instructional component. A study involving grade 2 classes has revealed striking differences between teachers regarding the amount of time they allow for play and the degree to which they stimulate pupils to engage in play. Generally speaking, one-third of classroom time is devoted to play and the remainder to

other activities. The differences between teachers with regard to play intervention also become evident in the classroom level of involvement in play. The degree to which teachers demonstrate effective classroom management has an even greater impact on involvement, however. Practical recommendations made within the framework of this project are: development of play programmes for younger children and in-service classroom management training for teachers.

Bijlage 1

Correlatiematrix van factoren en de variabele betrokkenheid van kinderen bij spelsituaties

	OPL	SES	ERV	ALG-ST	STUR	DCONT	ORG	BETR
OPL	-							
SES	0.00	-						
ERV	0.41	0.12	-					
ALGST	-.04	-.16	0.00	-				
STUR	-.02	-.24*	-.10	-.63*	-			
DCONT	0.13	-.25*	-.34	-.13	-.23*	-		
ORG	-.06	-.15	0.18	-.21*	-.12	0.41*	-	
BETR	-.01	0.17	0.12	0.20*	0.24*	-.31*	-.34*	-

Noot: * betekent significant bij $p<.05$

Teamgerichte nascholing en coaching: Effecten op het didactisch handelen van leerkrachten in combinatieklassen*

E. Roelofs, S. Veenman en J. Raemaekers**

Samenvatting

In dit onderzoek wordt verslag gedaan van de resultaten van de teamgerichte nascholing 'Omgaan met Combinatieklassen' zoals gegeven in het reguliere circuit van nascholing en begeleiding. Op vijf locaties in ons land verzorgden PABO-docenten de nascholing en werd coaching gegeven door schoolbegeleiders. In totaal namen 89 leerkrachten van 12 scholen aan deze nascholing deel, waarvan 10 scholen met het volledige team.

Onderzocht werden de effecten van het teamgerichte nascholingsprogramma en coaching op het didactisch handelen van de leerkrachten en de taakgerichtheid van de leerlingen. Daartoe werden voorafgaande aan en na afloop van de cursus observaties verricht in 42 klassen (experimentele groep met coaching n=18; zonder coaching n=10; controlegroep n=14). Uit de resultaten blijkt een duidelijk positief cursuseffect op de management- en instructievaardigheden en op de niveaus van taakgerichtheid. Daarenboven blijken gecoachte leerkrachten op enkele aspecten meer te profiteren van het nascholingsprogramma dan hun ongecoachte collega's. Tevens zijn er aanwijzingen gevonden dat effecten van coaching samenhangen met de mate van diepgang in de door begeleiders geboden hulp.

1 Inleiding

Steeds meer wordt ingezien dat nascholing in het kader van onderwijsvernieuwing zich niet primair moet richten op individuele leerkrachten, maar in de eerste plaats op het team van een school (ARBO, 1984, p. 56). Een belangrijk kenmerk van teamgerichte nascholing is de gerichtheid op de ontwikkeling van het schoolteam als georganiseerde (onderwijs)groep. Teamgerichte nascholing is daarmee te onderscheiden van individu-gerichte nascholing, welke in eerste instantie is gericht op de individuele leerkracht. Als belangrijk nadeel van individu-gerichte nascholing wordt naar voren gebracht dat deze weinig of geen veranderingen in de school als organisatie bewerkstelligt (Fullan, 1990).

2 Teamgerichte nascholing en coaching

Gedurende de laatste twee decennia zijn op basis van reviewstudies en synthesen van onderzoek naar vormgevingskenmerken van nascholing verschillende modellen voor teamgerichte nascholing ontworpen. Deze modellen hebben gemeen dat een belangrijk gewicht wordt toegekend aan componenten van effectieve training (Sparks & Loucks-Horsley, 1990). Het model met de sterkste empirische basis is dat van Joyce en Showers (1980, 1988). Het bevat de volgende componenten voor een effectieve training: 1) studie van de theorie die aan de te leren vaardigheden ten grondslag ligt, 2) observeren van een aantal demonstraties van de vaardigheden, 3) oefening in en toepassing van de vaardigheden, 4) terugkoppeling over de toepassing van de vaardigheden in gesimuleerde condities en in de eigen klas, en 5) coaching. In dit model wordt gestreefd naar een maximale transfer van vaardigheden in de klas. Modellen van Stallings (1985) en Sparks

* Het onderzoek is uitgevoerd met financiële steun van het Instituut voor Onderzoek van het Onderwijs (SVO-project 9343).

** De auteurs danken voor hun bijdrage aan het onderzoek: M. Voeten, M. Stadelman, N. Hofs, C. Evers, M. van Mierlo, E. Riemeijer en alle schoolteams, PABO-docenten en schoolbegeleiders die aan het onderzoek hebben deelgenomen.

(1983a,b) vertonen grote gelijkenis met dit model, met dit verschil dat meer nadruk gelegd wordt op onderzoek van en reflectie op het lesgedrag door leerkrachten. Bij teamgerichte nascholing wordt het belang van de school als organisatie benadrukt. Sparks (1983a,b) definieert teamgerichte nascholing als een 'nested process, that includes goals and content, the training process and the organizational context'. Joyce en Showers (1988) spreken van een systeem 'that pervades the school and is the responsibility of the organization'. Ondanks het belang dat gehecht wordt aan de school als organisatie zien we pas de laatste jaren dat verbetering van onderwijs in de klas en verbetering van de school als organisatie modelmatig worden geïntegreerd. Het model van schoolverbetering van Fullan (1990) is hiervan een voorbeeld. Nascholing wordt geflankeerd door veranderingen in de werkplek van leerkrachten en veranderingen in de schoolorganisatie.

Studies naar de relatie tussen vormgeving en effecten van nascholing laten zien dat nascholing pas leidt tot transfer, tot veranderingen in het didactisch handelen in de klas, wanneer naast training buiten de klas ook sprake is van begeleiding in de klas, ofwel coaching (Bennett, 1987). Bij programma's zonder coaching zijn transfereffecten minimaal.

Coaching is te beschouwen als een transferbevorderende interventie. Joyce en Showers (1980) omschrijven coaching als: 'Hands-on, in-classroom assistance with transfer and application of skills to the classroom' (p. 380). Coaching is een activiteit als vervolg op scholing of training waarbij de leerkracht ondersteuning ontvangt bij het toepassen van de nieuw geleerde inzichten en vaardigheden in de klas. Deze ondersteuning kan gegeven worden door een collega-leerkracht, de schoolleider of de schoolbegeleider.

Aan coaching kunnen vier functies worden toegekend (Joyce & Showers, 1988). De eerste functie betreft het geven van steun en vertrouwen en verwijst naar de mogelijkheid om via uitwisseling van ideeën na te denken over elkaars inzichten, te reflecteren op het eigen handelen en elkaars problemen, frustraties en successen te delen. De tweede functie betreft het geven van inhoudelijke terugkoppeling gericht op de geleerde vaardigheden. Coach en leerkracht richten hun aandacht op eventuele omissies in de toepassing van de vaardigheid. De gezamenlijke reflectie dient de perfectionering van de verworven vaardigheid. De derde functie van coaching, analyse van de toepassing, verwijst naar activiteiten van coach en leerkracht die erop gericht zijn de verworven vaardigheden op het juiste moment toe te passen. Aandacht wordt besteed aan de inpassing van vaardigheden bij het gebruik van leermiddelen, zoals de reken- of taalmethode. Geleerd wordt wannéér de vaardigheden het best gebruikt kunnen worden. Men kan hier ook spreken van contextualisering en decontextualisering van het geleerde (Simons, 1990). Een vierde functie betreft het afstemmen van het geleerde op de leerlingen. Nieuwe vaardigheden van de leerkracht kunnen voor leerlingen problematisch zijn. Daarom worden de reacties of gedragingen van de leerlingen met betrekking tot het onderwijsgedrag van de leerkracht geregistreerd en geanalyseerd. Deze functie is van belang, omdat de leerkracht in het begin meer met zichzelf en met de nieuwe vaardigheid bezig is dan met de leerlingen.

Meer specifiek blijken de volgende effecten op te treden bij gecoachte leerkrachten vergeleken met niet-gecoachte leerkrachten. Gecoachte leerkrachten blijken een helderder beeld te hebben van de nieuwe instructievaardigheden. Zij zijn beter in staat aan te geven wat het doel ervan is en welke mogelijkheden er zijn om de instructievaardigheden in verschillende situaties (contexten) toe te passen. Gecoachte leerkrachten passen nieuwe instructievaardigheden vaker toe en met een grotere deskundigheid. Zij passen de geleerde vaardigheden bovendien adequater toe dan niet-gecoachte leerkrachten. Niet-gecoachte leerkrachten neigen er toe precies datgene in praktijk te brengen wat tijdens de training is gedemonstreerd. Daarnaast blijken gecoachte leerkrachten de kennis en de vaardigheden langer vast te houden (Showers, 1984; Baker, 1983; Sparks, 1983a,b).

Nederlands onderzoek naar effecten van coaching, dat overigens schaars is, geeft een minder positief beeld. In een drietal experimentele studies vond Snippe (1988, 1991) dat coaching of klasseconsultatie nauwelijks (extra) effect had op de implementatie van een leergang cijferend rekenen.

3 Het nascholingsprogramma Omgaan met combinatieklassen

Vanaf 1981 werden aan de Vakgroep Onderwijskunde te Nijmegen enkele onderzoeken uitgevoerd naar het onderwijs in combinatieklassen. Elk volgend onderzoek bouwde daarbij voort op de resultaten van het voorgaande. In de fase van descriptief onderzoek (1981-1986) werd een aantal problemen van onderwijzen in combinatieklassen geïdentificeerd (Veenman, Lem, Voeten, Winkelmolen & Lassche, 1986). Deze problemen bleken gecentreerd te zijn rondom vijf gebieden: 1) het efficiënt gebruik van de taakgerichte leertijd, 2) effectieve instructie, 3) doelmatig klassemanagement, 4) het efficiënt inrichten van het zelfstandig werken en 5) de school- of teamgerichtheid met betrekking tot combinatieklassen. Op basis van deze vijf probleemgebieden werd in een tweede fase (start: 1986) een teamgerichte nascholingscursus ontwikkeld: 'Omgaan met combinatieklassen: een programma voor schoolverbetering' (Veenman, Lem & Nijssen, 1988). De inhoud van de cursus was gebaseerd op recente empirische inzichten over succesvol en effectief lesgeven. Voor een korte beschrijving van de cursusinhoud wordt verwezen naar een eerder artikel in dit tijdschrift (Veenman, Lem & Roelofs, 1990).

In het schooljaar 1986-1987 werd het programma in een experimentele setting voor de eerste maal beproefd. De resultaten uit het observatieonderzoek tijdens deze eerste experimentele beproeving waren zeer hoopgevend. De leerkrachten die hadden deelgenomen aan de cursus bleken allen een belangrijke vooruitgang te hebben geboekt wat betreft hun management- en instructiegedrag, hetgeen uitmondde in een hogere taakgerichtheid in hun klas. (Lem, Veenman, Nijssen & Roelofs, 1988; Veenman, Lem & Roelofs, 1990).

Punten van kritiek op de nascholing betroffen de relatief geringe toepasbaarheid van het programma voor leerkrachten van kleutergroepen. Om die reden is voor deze doelgroep een aanvullende cursuskatern ontwikkeld (Roelofs, Van Gennip & Veenman, 1989). Recentelijk is deze katern opgenomen in de nieuwste versies van het programma voor schoolverbetering (Veenman, Lem, Roelofs & Nijssen, 1993).

Een ander probleem tijdens de eerste beproeving was dat leerkrachten en schoolbegeleiders nauwelijks in staat waren om op een systematische wijze coaching te verzorgen. Er is al op gewezen dat coaching een belangrijk middel kan zijn om de transfer van verworven vaardigheden in de klas te bevorderen. In de tweede beproeving van het programma, waarvan in dit artikel verslag wordt gedaan, is op een meer systematische wijze coaching gegeven door schoolbegeleiders, die hiervoor een specifieke training ontvingen.

Een belangrijke vraag bleef echter of en in welke mate het nascholingsprogramma effectief was, wanneer het geplaatst zou worden in de reguliere context van nascholing en begeleiding. In onderhavige studie is het programma verzorgd door PABO-docenten, al dan niet in samenwerking met schoolbegeleiders, die het programma hebben aangeboden in hun reguliere nascholingsaanbod. Aan het bezwaar van de eerste evaluatie-studie, dat de rollen van cursusleiders en evaluatoren met elkaar verstrengeld waren, werd daarmee tegemoet gekomen.

4 Onderzoeksvragen

In deze studie wordt nagegaan of het teamgerichte nascholingsprogramma 'Omgaan met Combinatieklassen' (OMC) leidt tot verbeteringen in het didactisch handelen van leerkrachten, wanneer het wordt gegeven in de reguliere context van nascholing en begeleiding. Bovendien wordt nagegaan of de taakgerichtheid van de leerlingen hierdoor wordt beïnvloed. Daarnaast wordt onderzocht of er (extra) effecten te constateren zijn van coaching en of deze samenhangen met de vormgeving van coaching. Dit alles leidt tot de volgende algemene vraagstellingen:

1. Leidt het teamgerichte nascholingsprogramma 'Omgaan met combinatieklassen' (OMC) binnen de reguliere context van nascholing tot verbeteringen in management- en instructievaardigheden en daarmee tot een andere besteding van instructietijd door de leerkracht enerzijds en een hogere taakgerichtheid van leerlingen anderzijds? Wat

is de relatieve bijdrage van coaching aan eventuele effecten?
2. Is er een verband tussen de effecten van het programma OMC en de wijze waarop de coaching is gerealiseerd in de ogen van gecoachte cursisten?

Ook is het effect van het nascholingsprogramma op de leerresultaten van de leerlingen nagegaan. Hierover zal in een aparte publikatie worden gerapporteerd.

5 Opzet van het onderzoek

5.1 Vorming van de onderzoeksgroepen

Bij de vorming van de onderzoeksgroepen is zoveel mogelijk gepoogd rekening te houden met de gevaren van zelfselectie door de werving van de experimentele groep en de controlegroep over te laten aan de onderwijsbegeleidingsdiensten. Door de diensten is gericht gezocht naar schoolteams die behoefte hadden aan nascholing op het terrein van doelmatig klassemanagement en effectieve instructie. Door toewijzing van gelijksoortige schoolteams aan de experimentele en controleconditie konden op twee cursuslocaties (A en B) een vergelijkbare experimentele en controlegroep worden samengesteld. Op twee andere locaties (C en D) namen de schoolteams alleen deel aan de experimentele groep. Ter vergroting van de onderzoeksgroep werd contact opgenomen met een locatie waar de onderwijsbegeleidingsdienst het programma OMC reeds zelfstandig had gepland (locatie E). Men was bereid het programma te verschuiven om deel te kunnen nemen aan het onderzoek. De toewijzing van leerkrachten aan de condities 'coaching' en 'geen coaching' kwam tot stand door overleg tussen de schoolbegeleider en de schoolleider. Om te voorkomen dat scholen zouden afzien van deelname werden leerkrachten niet op basis van toeval of matching toegewezen aan deze condities. Dit bezwaar woog zwaarder dan het gevaar van zelfselectie.

5.2 Beschrijving van de onderzoeksgroepen

De teamgerichte nascholingscursus werd op vijf locaties uitgevoerd (Helmond, Nijmegen, Roosendaal, Sittard en Venlo). In totaal namen 89 leerkrachten deel die afkomstig waren van 12 scholen, waarvan 10 scholen met een voltallig team. De totale cursusgroep van 89 leerkrachten was als volgt samengesteld: 22 leerkrachten uit groep 1/2, 47 leerkrachten uit groep 3 t/m 8, 10 vaste inval-leerkrachten en 10 schoolleiders waarvan er twee gedeeltelijk ambulant waren.

Alle leerkrachten verleenden hun medewerking aan de evaluatie van de nascholing door middel van vragenlijsten. Leerkrachten van combinatieklassen (groep 3 t/m 8) verleenden medewerking aan de lesobservaties.

In Tabel 1 is de verdeling van de geobserveerde leerkrachten uit de experimentele groep over locaties en onderzoekscondities weergegeven. Uit de cursusgroep werden alle leerkrachten met een eigen combinatieklas (n=28) geobserveerd. Om praktische en budgettaire redenen werden de resterende leerkrachten, die in enkelvoudige klassen lesgaven, niet geobserveerd. Daarnaast waren alle 14 leerkrachten met combinatieklassen afkomstig van 6 controlescholen bereid observaties toe te staan in hun klassen. Deze controlegroep bestond uit 4 scholen uit locatie A (per school werden 1 à 3 leerkrachten geobserveerd, n=8) en 2 scholen uit locatie B (per school werden hier 3 leerkrachten geobserveerd, n=6). Het totaal aantal geobserveerde leerkrachten in de experimentele groep en de controlegroep samen bedroeg hiermee 42.

De samenstelling van de experimentele groep en de controlegroep verschilde niet sterk van elkaar: de gemiddelde leeftijd bedroeg resp. 40.4 (sd=6.7) en 42.2 jaar (sd=8.5), het aantal jaren onderwijservaring resp. 18.4 jaar (sd=6.3) en 20.0 jaar (sd=8.5). In de experimentele groep was 71% van de leerkrachten een man, in de controlegroep 64%. De samenstelling van de groep gecoachte en niet-gecoachte leerkrachten verschilde qua leeftijdssamenstelling: de gecoachte leerkrachten bleken iets ouder te zijn (gem: 42.1 jaar, sd= 5.5) dan de niet-gecoachte leerkrachten (gem: 36.9 jaar, sd=8.0). Daarmee gaat tevens een verschil in ervaring gepaard (gem: 19.9, sd=5.1 resp. gem: 13.3, sd=7.8).

Behalve verschillen in samenstelling is getoetst of de onderzoeksgroepen op de voormeting van elkaar verschilden voor wat betreft de criteriumvariabelen van het onderzoek. De totale experimentele groep en de controlegroep

Tabel 1
Verdeling van de geobserveerde leerkrachten uit de experimentele groep over de locaties en onderzoekscondities

Locatie	Geobserveerde leerkrachten (combinatieklas, groep 3 t/m 8)		
	Aantal cursisten/omvang van het schoolteam	Gecoacht	Niet-gecoacht
Locatie A:			
School 1	7	3	-
School 2	5	-	1
School 3	6	-	1
Locatie B:			
School 4	5	2	2
School 5	8	2	1
School 6	8	3	-
Locatie C:			
School 7	8	1	2
School 8	10	2	-
Locatie D-1:			
School 9	18	3	-
Locatie D-2[1]:			
Scholen 10/11	4	-	3
Locatie E:			
School 12	10	2	-
Totaal	89	18	10

[1]: Op locatie D-2 namen 4 leerkrachten van 2 scholen op individuele basis aan de nascholing deel. Van de 89 cursisten ontvingen 36 leerkrachten coaching, waarvan de helft aan de observatiestudie heeft deelgenomen.

verschilden niet significant van elkaar wanneer gekeken werd naar management- en instructievaardigheden van de leerkrachten en de besteding van leer- en instructietijd door leerkracht en leerlingen. Wel bleken niet-gecoachte leerkrachten hoger te scoren op een aantal managementvaardigheden dan gecoachte leerkrachten. Met dit gegeven dient rekening te worden gehouden bij de interpretatie van de resultaten.

5.3 Inrichting van de teamgerichte nascholing en de coaching

Ter voorbereiding op de daadwerkelijke nascholing en coaching werd ten behoeve van de betrokken PABO-docenten en schoolbegeleiders een voorscholing verzorgd. De deelnemers gaven bij de evaluatie aan dat deze voorscholing in redelijke mate had bijgedragen aan de eigen professionalisering met het oog op de uitvoering van de teamgerichte nascholingscursus.

Het nascholingsprogramma werd verzorgd door ervaren PABO-docenten. Op één locatie verzorgde een schoolbegeleider zowel de cursus als de coaching. De cursus 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. Tussen de bijeenkomsten in kreeg bovendien een deel van de leerkrachten coaching, verzorgd door de schoolbegeleider en in een enkel geval door de schoolleider.

5.4 Instrumentarium

Met behulp van het observatie-instrument COMMIT-II werd een schatting verkregen van de tijd die door de leerlingen en leerkrachten in combinatieklassen aan verschillende activiteiten en settings wordt besteed. De observatoren registreerden: 1) de taakgerichte leertijd van leerlingen, 2) de groeperingsvorm (setting) waarin de leerlingen werkten, 3) de activiteit van de leerkracht en 4) de doelgroep van de leerkracht. Bij COMMIT-II wordt gebruik gemaakt van een specifieke vorm van 'time-sampling', namelijk de 'predominant activity sampling'-procedure (Tyler, 1979). De procedure verliep als volgt: eerst werden de activiteiten van leerling één en de leerkracht geobser-

veerd (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 gecodeerd, 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. Begonnen werd met leerling één van de laagste (jaar)groep. De leerkracht werd niet op de hoogte gebracht van de volgorde om beïnvloeding van het lesgedrag te voorkomen.

Met behulp van de Management- en Instructie Schaal (MIS) werden meer specifieke gegevens verzameld over de kwaliteit van de management- en instructievaardigheden van de leerkracht. De MIS is gebaseerd op onderzoek naar effectief managementgedrag van leerkrachten (Evertson, Emmer, Clement, Sanford & Worsham, 1984; Veenman, Lem, Voeten, Winkelmolen & Lassche, 1986) en op onderzoek naar kenmerken van effectieve instructie (Rosenshine & Stevens, 1986; Good, Grouws & Ebmeier, 1983). De 41 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. De subschaal 'instructievaardigheden' werd ten behoeve van dit onderzoek ten opzichte van de vorige versie uitgebreid met 10 items om een betere representatie te verkrijgen van het model van directe instructie (zie ook Roelofs, 1993).

Na een uitgebreide training van 80 uur voorafgaand aan de voormeting bleek de tussen-observator-betrouwbaarheid van de COMMIT-II goed te zijn: deze varieerde voor de verschillende observatie-categorieën van .65 tot 1.0 (gemiddelde: .85). Het zelfde gold voor de tussen-observator-betrouwbaarheid van de subschalen van de MIS: deze liep uiteen van .67 tot .92 (gemiddelde: .83). Vlak voor de nameting van het onderzoek werd een opfristraining gegeven. De inter-observator-betrouwbaarheid handhaafde zich op een zelfde niveau.

Om de vormgeving van de coaching door de cursisten te laten beoordelen, werd een vragenlijst ontwikkeld. Bij de ontwikkeling van deze vragenlijst werd gelet op de vormgeving van relevante functies van coaching zoals onderscheiden door Joyce en Showers (1980, 1988). De volgende aspecten kwamen aan de orde: professionele steun, terugkoppeling in algemene zin en specifieke zin, inpassing van cursusvaardigheden bij het gebruik van de methode (analyse van de toepassing), bruikbaarheid van de inpassing, afstemming van de vaardigheid op de leerlingen en aansluiting op de cursusinhoud. De lijst werd afgenomen bij de totale groep van gecoachte cursisten (n=36).

Voorts werd een vragenlijst afgenomen waarin de cursisten hun waardering over de nascholing konden uitspreken, alsmede een lijst waarin de cursisten werden bevraagd naar hun schoolorganisatiebeleving. De resultaten op de twee laatstgenoemde lijsten die aan alle cursisten werden voorgelegd (n=89, waarvan 76 retour) worden in dit artikel slechts zijdelings behandeld. Voor meer gedetailleerde informatie wordt verwezen naar Roelofs (1993).

5.5 Opzet en analyse

De toetsing van de effecten van de nascholing en coaching geschiedde door middel van een 'non-equivalent control group design' met voor- en nameting. Het design is schematisch weergegeven in Tabel 2. Het onderzoek kende twee experimentele condities (treatments) en één controleconditie. De leerkrachten in de experimentele conditie zónder coaching ontvingen alleen nascholing, welke bestond uit een cursus van vijf tot zeven bijeenkomsten ver-

Tabel 2

Opzet van het onderzoek voor het bepalen van effecten van de teamgerichte nascholingscursus en coaching

	n Lk	Y1	X1	X2	Y2
Experimentele groep met coaching	18	+	+	+	+
Experimentele groep zonder coaching	10	+	+	-	+
Controlegroep	14	+	-	-	+

N.B. *n* lk= aantal leerkrachten Y: observaties gedurende vier lessen; Y1= voormeting, Y2=nameting, X1= nascholing, X2= coaching.

zorgd door een PABO-docent eventueel in samenwerking met een schoolbegeleider. De nascholing vond op alle cursuslocaties plaats in de periode van 15 januari 1990 tot 1 april 1990. De leerkrachten in de experimentele conditie mét coaching ontvingen behalve nascholing ook coaching. De coaching bestond uit extra ondersteuning door de schoolbegeleider of de schoolleider in de klas gedurende de periode tussen twee nascholingsbijeenkomsten.

De leerkrachten uit de controlegroep ontvingen in de periode van januari tot april 1990 geen treatment. Zij werden in globale termen op de hoogte gebracht van het onderzoek. Uit contacten met deze leerkrachten tijdens het onderzoek bleek dat zij geen behoefte hadden aan een meer gedetailleerde toelichting.

Het aantal geobserveerde leerkrachten uit de conditie met coaching bedraagt 18, het aantal geobserveerde leerkrachten in de conditie zonder coaching 10. Daarnaast werden 14 leerkrachten uit de controlegroep geobserveerd. De leerkrachten uit alle condities werden zowel tijdens de voor- als de nameting gedurende twee lees-/taal- en twee rekenlessen geobserveerd. De voormeting had plaats in de periode van 2 tot 12 weken voorafgaand aan de cursus, de nameting in de periode van 2 tot 12 weken na afloop van de cursus. Alle observaties vonden plaats in de ochtend, waarbij één les voor de pauze en één les na de pauze werd geobserveerd. Dit betekende dat de observaties per leerkracht en per meting twee ochtenden in beslag namen.

Ter bescherming van de interne validiteit van de observatiegegevens werd een drietal maatregelen genomen. Ten eerste werden de observatie-ochtenden per leerkracht gespreid met een tussenperiode van minimaal twee weken. Ten tweede werd er voor gezorgd dat in elke observatie-week leerkrachten uit alle drie condities werden geobserveerd. Ten derde werd elke leerkracht door zoveel mogelijk verschillende observatoren geobserveerd.

De observatiegegevens verkregen met behulp van COMMIT-II hadden per meting steeds betrekking op de gemiddelde tijdsbesteding van de hele klas, berekend over de vier geobserveerde lessen. Per observatiecategorie resulteerde een tijdsschatting die werd uitgedrukt in minuten. De observatiegegevens van de MIS werden omgezet in somscores (berekend door de itemscores binnen elk van de vijf subschalen te sommeren). Deze scores hadden eveneens betrekking op de gemiddelden over vier lessen.

Effecten van de cursus op de tijdsbesteding en de management- en instructievaardigheden van de leerkracht werden getoetst door middel van een *t*-toets voor verschilscores. Hierbij werden de gemiddelde verschilscores van de totale experimentele groep afgezet tegen de gemiddelde verschilscores van de controlegroep. Voor de toetsing van effecten van coaching werden alleen de data geanalyseerd van cursisten die afkomstig waren van die cursuslocaties waar naast gecoachte ook niet-gecoachte leerkrachten voorkwamen. Deze leerkrachten konden worden vergeleken met collega's die deelnamen aan dezelfde cursus, maar geen coaching ontvingen, waardoor het effect van de coaching gescheiden kon worden van het cursuseffect. Immers, leerkrachten van een locatie met alleen nascholing konden niet zonder meer worden vergeleken met collega's van een locatie waar zowel nascholing als coaching werden gegeven. Hierdoor werd de experimentele groep met coaching teruggebracht tot 13 leerkrachten en de experimentele groep zonder coaching tot 7 leerkrachten.

De inhoud van de nascholing was omschreven in het schoolverbeteringsprogramma 'Omgaan met Combinatieklassen'. Omdat de uitvoering van de nascholing per cursuslocatie kon verschillen door het leggen van eigen accenten, werd een aanvullende analyse uitgevoerd op de observatiegegevens. Om die reden zijn de veranderingen in leerkracht- en leerlinggedrag nogmaals geanalyseerd, waarbij de locatieverschillen uit de errorterm zijn gehaald en als afzonderlijke factor in de analyses meegenomen. Gebruik is gemaakt van een variantie-analyse in een één-factorieel design met 10 groepen: één controlegroep en 9 experimentele groepen. De experimentele groepen zijn onderscheiden op basis van de locatie waar de cursus werd uitgevoerd (verschillen tussen cursusleiders) en het al dan niet toegepast zijn van coaching. Op drie locaties (A, B, C) werd een deel van de geobserveerde leerkrachten gecoacht en een deel niet, leidend tot zes groepen. Op twee andere locaties (D, E) werden alle geobserveerde leerkrachten gecoacht dan wel niet gecoacht, leidend tot drie groepen. Door

Tabel 3

Gemiddelde besteding van leer- en instructietijd in de experimentele groep en de controlegroep op de voor- en de nameting (in minuten[1]; lesduur: 40 minuten)

	Controlegroep (*n*=14)			Exp. Groep (*n*=28)			
Observatiecategorie	Voor *Gem*	Na *Gem*	Verschil	Voor *Gem*	Na *Gem*	Verschil	*p*
Activiteiten leerling:							
Taakgericht	27.7	28.5	0.7	27.9	30.4	2.5	*
Procedureel	2.9	2.8	-0.1	3.2	2.3	-1.0	*
Wachten	1.7	1.5	-0.1	1.9	1.3	-0.5	n.s
Niet met de opdracht bezig	7.8	7.2	-0.5	7.0	6.1	-1.0	n.s
Setting en taakgerichtheid (in percentages lestijd):							
Taakgerichtheid tijdens instructie	75.7	80.6	4.9	72.0	80.1	8.1	n.s
Taakgerichtheid tijdens verwerking	65.8	66.1	0.3	67.7	74.1	6.4	*
Leerkrachtactiviteiten:							
Instructie (totaal)	22.8	20.9	-1.9	21.8	22.7	0.9	n.g
Terugblik	0.1	0.4	0.2	0.2	1.6	1.5	**
Uitleg	21.6	20.5	-1.1	21.0	20.1	-0.9	n.g
Geleide oefening	1.0	0.1	-1.0	0.6	1.0	0.4	**
Begeleiding:							
Hulp	7.9	7.6	-0.3	6.3	6.0	-0.3	n.g
Gerichte controle	4.0	6.2	2.3	5.0	5.1	0.0	*‡
Management:							
Lesovergangen/organisatie	3.9	3.8	-0.1	5.4	5.2	-0.2	n.s
Voorbereidingen/eigen werk	1.5	1.5	0.0	1.5	1.1	-0.4	n.s

[1]: = behalve de categorie 'setting en taakgerichtheid'; n.g= niet getoetst op implementatie omdat toename van dit instructiegedrag geen oogmerk vormde van de nascholing; n.s= niet significant; *= $p<.05$ **=$p<.01$; ‡= tegengesteld aan de verwachting.

Tabel 4

Gemiddelde besteding van leer- en instructietijd in de experimentele groepen met en zonder coaching tijdens de voor- en de nameting (in minuten[1]; lesduur: 40 minuten)

	Exp. met coaching (*n*=13)			Exp. zonder coaching (*n*=7)			
Observatiecategorie	Voor *Gem*	Na *Gem*	Verschil	Voor *Gem*	Na *Gem*	Verschil	*p*
Activiteiten leerling:							
Taakgericht	26.8	30.2	3.3	29.4	30.2	0.8	*
Procedureel	3.2	2.0	-1.2	3.0	2.7	-0.3	n.s
Wachten	2.0	1.4	-0.5	1.6	1.3	-0.3	n.s
Niet met de opdracht bezig	8.0	6.4	-1.6	6.0	5.8	-0.2	n.s
Setting en taakgerichtheid (in percentages lestijd):							
Taakgerichtheid tijdens instructie	70.0	79.8	9.8	75.3	79.7	4.4	n.s
Taakgerichtheid tijdens verwerking	64.6	73.9	9.3	71.4	72.4	1.0	*
Leerkrachtactiviteiten:							
Instructie (totaal)	23.8	24.2	0.4	19.5	20.4	0.9	n.g
Terugblik	0.2	1.2	1.0	0.1	1.6	1.4	n.s
Uitleg	22.7	21.8	-0.9	19.1	18.0	-1.1	n.g
Geleide oefening	0.9	1.1	0.3	0.2	0.9	0.6	n.s
Begeleiding:							
Hulp	4.8	4.8	0.0	10.1	7.3	-2.8	n.g
Gerichte controle	4.1	4.8	0.7	4.7	5.4	0.7	n.s
Management:							
Lesovergangen/organisatie	5.8	5.2	-0.6	3.9	5.3	1.4	*
Voorbereidingen/eigen werk	1.6	0.9	-0.6	1.8	1.6	-0.2	n.s

[1]: = behalve de categorie 'setting en taakgerichtheid', hier is sprake van %; n.g= niet getoetst op implementatie omdat toename van dit instructiegedrag geen oogmerk vormde van de nascholing; n.s= niet significant; *= $p<.05$.

deze aanpak werden locatieverschillen uit de errorterm gehaald, waardoor significantietoetsingen gevoeliger zijn. Waar deze analyse leidde tot andere resultaten, wordt dit vermeld.

6 Resultaten

In de Tabellen 3 en 4 wordt een overzicht gegeven van de gemiddelde besteding van leer- en instructietijd per onderzoeksgroep vóór en na de nascholingscursus, alsmede verschillen tussen de voor- en de nameting.

De resultaten van Tabel 3 wijzen uit dat het niveau van taakgerichtheid in de experimentele groep is gestegen (30.4 minuten, 76% van de lestijd) ten opzichte van de voormeting (27.9 minuten, 70% van de lestijd). Dit verschil is significant wanneer dit wordt vergeleken met de veranderingen in taakgerichtheid in de controlegroep (toename: resp. 0.7 en 2.5 minuut, $p < .05$, zie Tabel 3). De stijging in taakgericht gedrag is gepaard gegaan met een daling in tijd besteed aan procedurele activiteiten. De tijd besteed aan wachten en niet-met-de-opdracht-bezig-zijn neemt weliswaar af, maar deze laatste verschillen zijn niet significant ten opzichte van de controlegroep.

Uit de resultaten van Tabel 4 blijkt dat het niveau van taakgericht gedrag in klassen van gecoachte leerkrachten sterker is gestegen (toename 3.4 minuten of 9%) dan in klassen van niet-gecoachte leerkrachten (toename 0.8 minuten of 2%). Deze extra winst in taakgerichte leertijd kan niet toegeschreven worden aan één bepaalde bron van niet-taakgericht gedrag. Zowel de hoeveelheid tijd besteed aan procedurele activiteiten, als die aan wachten en niet-met-de-opdracht-bezig-zijn, is sterker afgenomen voor de gecoachte dan de niet-gecoachte groep, maar deze verschillen zijn niet significant.

Omdat de duur van een bepaalde setting per klas kan variëren, is de taakgerichtheid per setting (de taaksituatie waarin de leerlingen verkeren) uitgedrukt in percentages taakgerichte leertijd per setting. De nascholingscursus heeft geleid tot een toename in taakgericht gedrag tijdens instructie en verwerking. De taakgerichtheid tijdens instructie neemt in de experimentele groep weliswaar toe (80.1% ten opzichte van 72.0%) maar bijna eenzelfde toename is ook zichtbaar in de controlegroep (80.6% ten opzichte van 75.7%). Hierdoor is geen sprake van een significant cursuseffect (zie Tabel 3). Op de voormeting bedroeg het percentage taakgerichtheid tijdens individuele verwerking 67.7, op de nameting bedroeg dit percentage 74.1. Deze toename is significant hoger dan die in de controlegroep. De toename in taakgerichtheid tijdens verwerking is bij gecoachte leerkrachten significant groter (toename: 9.3%) dan in klassen van niet-gecoachte leerkrachten (toename: 1.0%). In beide groepen is ook sprake van een toename van taakgerichtheid tijdens instructie. Hoewel deze toename in de gecoachte groep (toename: 9.8%) groter is dan in de niet gecoachte groep (toename: 4.4%), is geen sprake van een significant verschil tussen beide groepen.

De besteding van lestijd door leerkrachten kan gesplitst worden in besteding van lestijd aan instructie, aan activiteiten tijdens verwerking (begeleidingsactiviteiten) en aan activiteiten met betrekking tot lesorganisatie (managementactiviteiten).

De hoeveelheid tijd die besteed wordt aan instructie (zie Tabel 3) is in de experimentele groep op de nameting relatief weinig veranderd (22.7 minuten) ten opzichte van de voormeting (21.8 minuten). In de wijze waarop de instructietijd wordt besteed zijn wel enkele verschuivingen opgetreden in de verwachte richting. Leerkrachten die hebben deelgenomen aan de cursus besteden tijdens de nameting meer tijd aan het geven van een terugblik (1.6 minuut ten opzichte van 0.2 minuut op de voormeting) en aan het geleid inoefenen van de leerstof (1.0 minuut ten opzichte van 0.6 minuut). Beide verschuivingen vormen significante cursuseffecten.

Voor wat betreft de tijdsbesteding aan de instructie-onderdelen terugblik en geleide oefening zijn geen effecten van coaching te constateren. De toename in tijdsbesteding aan deze onderdelen is gelijk voor wel en niet-gecoachte leerkrachten.

Ten aanzien van gerichte controle, een activiteit ten behoeve van de begeleiding van leerlingen tijdens individuele verwerking, kan niet gesproken worden van een positief significant cursuseffect (zie Tabel 3). Leerkrachten uit de experimentele groep besteden na afloop van de cursus niet meer tijd aan gerichte controle op

Tabel 5

Scores op de Management- en InstructieSchaal van de experimentele groep en de controlegroep tijdens de voor- en de nameting

Subschaal MIS	Controlegroep (*n*=14) Voor *Gem*	Na *Gem*	Verschil	Exp. Groep (*n*=28) Voor *Gem*	Na *Gem*	Verschil	*p*
Instructievaardigheden (17 items: α= .87)	56.4	56.8	0.4	54.3	63.2	8.9	**
Organisatie van de instructie (5 items: α= .81)	19.5	19.1	-0.4	18.2	20.1	1.9	**
Gebruik van materiaal en ruimte (6 items: α= .81)	25.4	26.0	0.6	24.5	27.0	2.4	*
Afstemming van de instructie (6 items: α= .91)	22.9	22.1	-0.9	21.9	23.6	1.7	**
Omgaan met storingen (6 items: α= .91)	20.9	21.9	1.0	20.2	23.8	3.6	*

N.B. De gemiddelden hebben betrekking op somscores over de samenstellende items; *= $p<.05$ **=$p<.01$.

Tabel 6

Scores op de Management- en InstructieSchaal van de experimentele groepen met en zonder coaching tijdens de voor- en de nameting

Observatiecategorie	Exp. met coaching (*n*=13) Voor *Gem*	Na *Gem*	Verschil	Exp. zonder coaching (*n*=7) Voor *Gem*	Na *Gem*	Verschil	*p*
Instructievaardigheden (17 items: α= .87)	53.8	61.2	7.4	56.0	64.4	8.4	n.s
Organisatie van de instructie (5 items: α= .81)	17.5	19.7	2.2	19.9	20.5	0.6	*
Gebruik van materiaal en ruimte (6 items: α= .81)	23.9	26.7	2.8	24.8	27.5	2.7	n.s
Afstemming van de instructie (6 items: α= .91)	21.0	22.7	1.8	23.9	24.7	0.8	n.s
Omgaan met storingen (6 items: α= .91)	18.6	23.0	4.4	23.1	25.2	2.0	*

N.B. De gemiddelden hebben betrekking op somscores over de samenstellende items; n.s= niet significant; *= $p<.05$.

het gedrag of werkzaamheden van de leerlingen. Opvallend is juist dat leerkrachten uit de controlegroep significant meer gerichte controle uitoefenen (toename: 2.2 minuten) vergeleken met hun collega's uit de experimentele groep (praktisch gelijk gebleven). Er zijn geen verschillen tussen gecoachte en niet-gecoachte leerkrachten ten aanzien van verandering in tijd besteed aan gerichte controle.

De verwachte afname in tijd besteed aan managementgedrag, waaronder leswisselingen en andere organisatorische activiteiten, treedt niet op in de totale experimentele groep (zie Tabel 3). Deze tijd neemt niet significant af ten opzichte van de voormeting (5.2 minuten ten opzichte van 5.4 minuten). De tijd waarin leerkrachten uit de experimentele groep bezig zijn met het treffen van voorbereidingen op nieuwe lesactiviteiten en het verrichten van eigen werkzaamheden, neemt wel af (afname: 0.4 minuten), maar niet significant ten opzichte van de controlegroep (geen verschil). Er is wel een effect van coaching zichtbaar op de tijd besteed aan organisatie en lesovergangen, dat bestaat uit enerzijds een toename van deze tijd in de groep niet-gecoachte leerkrachten en anderzijds een afname van deze tijd in de groep gecoachte leerkrachten.

In de Tabellen 5 en 6 zijn de resultaten op de MIS op de voor- en de nameting weergegeven, alsmede de resultaten van de *t*-toetsen op de verschilscores tussen voor-en nameting. Uit Tabel 5 blijkt dat de instructievaardigheden van leerkrachten uit de experimentele groep duidelijk zijn verbeterd. Het verschil ten opzichte van de controlegroep is significant. In het algemeen betekent dit dat deze leerkrachten op de nameting in hogere mate gebruik maken van het directe instructiemodel met daarin de componenten: terugblik, presentatie, geleide oefening, individuele verwerking en regelmatige terugkoppeling. De mate van de afstemming van instructie op de te bereiken doelen en op het niveau van de leerlingen is ook signifi-

cant verbeterd. Verder zijn cursuseffecten te constateren op de kwaliteit van de organisatie van de instructie. Hiermee wordt bedoeld het geheel aan voorwaarden en maatregelen om een goede instructie mogelijk te maken. De managementvaardigheden die betrekking hebben op het gebruik van materiaal en ruimte en op de omgang met storingen kennen een duidelijke verbetering in de cursusgroep. Kortom, de cursus heeft op alle onderscheiden management- en instructievaardigheden een positief effect gehad.

Bij vergelijking van de gecoachte en niet-gecoachte groep leerkrachten blijkt dat op twee van de vijf subschalen van de MIS een grotere winst zichtbaar is in de gecoachte groep cursisten. Op de subschalen 'organisatie van de instructie' en 'omgang met storingen' boeken gecoachte leerkrachten significant meer vooruitgang dan hun niet-gecoachte collega's.

De analyses waarin wordt gecorrigeerd voor locatieverschillen leiden tot een bevestiging van de gevonden resultaten. Om die reden worden hierover geen gegevens gepresenteerd (zie Roelofs, 1993). Vermeldenswaardig is de bevinding dat de effecten van nascholing en coaching op de activiteiten van de leerlingen zelfs wat groter zijn dan bij de toetsing zonder correctie voor locatieverschillen. Zo blijkt dat de wachttijd in klassen van cursisten sterker is afgenomen dan in klassen van de controlegroep, hetgeen in de voorgaande analyses niet kon worden aangetoond. Verder is de afname in tijd besteed aan procedurele activiteiten en aan niet-bij-de-les-betrokken gedrag sterker in klassen van gecoachte leerkrachten. Ook deze effecten waren in eerdere analyses niet significant. Voorts bleek een zeker patroon in de locatie-effecten te zitten: locaties verschilden significant van elkaar op het aspect van positief managementgedrag van de leerkracht, gerichte controle en organisatie van de instructie. Deze verschillen bleken vervolgens terug te vinden in locatieverschillen met betrekking tot het taakgerichte gedrag van de leerlingen.

Om de samenhang tussen vormgeving en effecten van coaching na te gaan, zijn achteraf via een clusteranalyse op de verschilscores op de subschaal 'instructievaardigheden' (nameting minus voormeting) volgens de WARD-methode twee groepen leerkrachten gevormd: zij die resp. in sterke en geringe mate vooruit zijn gegaan op deze subschaal. De gemiddelde verschilscores bedroegen respectievelijk 12.6 (*sd*=2.6) en 4.9 (*sd*=2.6). Gekozen is voor deze subschaal, omdat deze betrekking heeft op de vaardigheden die het meest centraal staan in de cursus. Overigens blijkt de tweedeling ook terug te vinden in de scores op de overige subschalen van de MIS. In Tabel 7 is een aantal belangrijke vormgevingsaspecten van coaching weergegeven, zoals deze zijn beoordeeld met behulp van de daartoe ontwikkelde vragenlijst 'Evaluatie van de ontvangen coaching'. De schaalwaarde van elk item loopt van 1 ('geheel niet') tot 5 ('in zeer ruime mate'). Voor de verantwoording van de vragenlijst en de vorming van subschalen wordt verwezen naar Roelofs (1993).

Uit Tabel 7 blijkt dat de coaching in het algemeen positiever gewaardeerd wordt door leerkrachten die sterk zijn vooruitgegaan met betrekking tot instructievaardigheden. Vergeleken met leerkrachten uit de lage effectgroep oordelen zij positiever over de mate waarin sprake was van een uitwisseling van suggesties en ideeën. De coach was naar hun mening actiever in het aanreiken van concrete suggesties en ideeën ten behoeve van de lesuitvoering. De twee effectgroepen verschillen nauwelijks van elkaar in de beoordeling van de gegeven terugkoppeling over het eigen lesgedrag. Opvallend is wel dat juist bij leerkrachten in de lage effectgroepen veel meer gewerkt is aan de hand van feedbacklijsten.

Over de analyse van de toepassing, ofwel de inpassing van cursusvaardigheden in de dagelijkse lespraktijk, zijn duidelijke verschillen in waardering te constateren tussen de twee effectgroepen (zie Tabel 7). In de hoge effectgroep is met name sprake geweest van een betere hulp bij de inpassing van cursusinhouden in de dagelijkse lespraktijk. De leerkrachten in deze groep geven aan meer tevreden te zijn over de wijze waarop getracht is de cursusvaardigheden in te passen bij de gebruikte rekenmethode. Daarnaast zijn ze in hogere mate tevreden over de praktische haalbaarheid van de aangedragen ideeën ten aanzien van de inpassing in de praktijk vergeleken met leerkrachten uit de lage effectgroep.

Wat betreft de aanpassing van de onderwijsaanpak aan de leerlingen, volgen de verschillen

Tabel 7

Waardering van de coaching door leerkrachten met een sterke en een geringe vooruitgang op instructievaardigheden; toetsing van verschillen door middel van t-tests

VORMGEVINGSKENMERKEN VAN COACHING	Hoog effect (*n*=6)		Laag effect (*n*=9)		
	Gem	*sd*	*Gem*	*sd*	*p*
Professionele steun (3 items: α=.80)	4.0	0.4	3.3	0.6	n.s.
1. De coach wisselde ideeën/suggesties uit met de leerkracht	4.0	0.6	3.2	0.8	*
2. De coach reikte suggesties en ideeën aan	4.2	0.4	3.4	0.5	*
3. De relatie tussen coach en leerkracht heeft bijgedragen aan de implementatie van cursusvaardigheden	3.8	0.7	3.4	0.7	n.s.
Globale terugkoppeling (4 items: α=.87)	3.9	0.3	3.7	0.6	n.s.
1. De coach besprak punten ter verbetering in leerkrachtgedrag	4.0	0.0	3.9	0.6	n.s.
2. De gegeven terugkoppeling was bruikbaar bij de planning van volgende lessen	4.0	0.0	3.5	1.1	n.s.
3. De coach hanteerde feedbacklijsten	2.6	1.1	3.9	0.6	*
4. De coach hanteerde aandachtspunten van de leerkracht	3.8	0.7	3.6	1.1	n.s.
Specifieke terugkoppeling (2 items: α= .65)	3.6	0.2	3.3	0.8	n.s.
1. De terugkoppeling werd gekoppeld aan cursusopdrachten	3.2	0.4	3.0	0.9	n.s.
2. De terugkoppeling werd gekoppeld aan de cursusinhoud	3.8	0.4	3.5	0.9	n.s.
Analyse van de toepassing (AT, 7 items: α=.83)	3.7	0.6	3.1	0.6	*
1. Experimenteren met nieuwe werkwijzen	3.3	0.8	3.4	0.7	n.s.
2. AT: het directe instructiemodel bij lees-/taalmethode	3.5	1.4	3.1	1.1	n.s.
3. AT: het directe instructiemodel bij de rekenmethode	4.6	0.5	3.3	0.7	**
4. AT: zelfstandig werken tijdens lees-/taallessen	3.3	1.4	3.0	0.7	n.s.
5. AT: zelfstandig werken tijdens de rekenlessen	4.6	0.5	2.9	0.8	**
6. Praktische haalbaarheid van suggesties en ideeën	3.7	0.8	2.8	0.7	*
7. Duidelijkheid van aangereikte suggesties en ideeën	3.5	1.0	3.1	0.6	n.s.
Aanpassing aan de leerlingen (3 items: α= .73)	2.8	1.1	2.4	0.5	n.s.
1. De coach wees op reacties van leerlingen	3.0	1.0	2.6	0.7	n.s.
2. De coach ontwikkelde een werkwijze om leerlingen te laten wennen aan de veranderde lesaanpak	2.6	1.3	2.1	0.6	n.s.
3. De coach hield rekening met verschillen tussen leerlingen	3.0	1.1	2.6	1.2	n.s.
Aansluiting bij de cursusinhoud (4 items: α= .73)	3.7	0.3	3.2	0.4	*
1. Aansluiting bij leertijd	3.2	1.0	3.4	0.7	n.s.
2. Aansluiting bij instructie	4.2	0.4	3.7	0.5	*
3. Aansluiting bij klassemanagement	3.5	0.8	2.9	0.6	n.s.
4. Aansluiting bij zelfstandig werken	4.0	0.6	2.8	0.9	*

N.B. * = $p < .05$, ** = $p < .01$.

de algemene trend, hoewel geen sprake is van significante verschillen. In de hoge effectgroep lijkt tijdens de coaching wat meer aandacht te zijn geweest voor de leerlingen (reacties van leerlingen en leerlingen laten wennen aan de nieuwe lesaanpak). De invulling van deze functie heeft overigens over het algemeen minder aandacht gekregen dan de overige functies.

Tot slot zijn de leerkrachten uit de hoge effectgroep meer tevreden over de aansluiting van de coaching bij de cursusinhoud dan hun collega's uit de lage effectgroep. Dit geldt met name voor de cursusonderwerpen instructie en zelfstandig werken.

De groepen leerkrachten verschillen niet van elkaar in de waardering van de cursus. Op geen van de onderscheiden schalen waarmee deze waardering is gemeten, zijn verschillen gevonden, met uitzondering van de waardering van de planning van de bijeenkomsten. Deze planning is in de ogen van de groep 'hoog effect' beter geweest dan in de groep 'laag effect'. Wat betreft schoolorganisatiebeleving zijn evenmin verschillen gevonden tussen leerkrachten uit de hoge en de lage effectgroep. Op de subschalen 'klimaat', 'doelgerichtheid' en 'gerichtheid op vernieuwingen' van de vragenlijst Schoolkenmerken (Van Tulder, Veenman

& Roelofs, 1989) scoren beide groepen ongeveer gelijk. Wel is er een tendens die erop wijst dat leerkrachten uit de hoge effectgroep de cursusinhoud meer in overeenstemming achten met hun eigen opvattingen over lesgeven (gem. 3.8 versus 3.4 op een 5-puntsschaal). Of dit positieve oordeel een voorwaarde danwel een gevolg is van de positieve waardering van de coaching kan op basis van deze gegevens niet worden uitgemaakt. Deze bevindingen indiceren al met al dat de gevonden onderlinge verschillen in vooruitgang tussen gecoachte leerkrachten samenhangen met de wijze waarop de coaching is gerealiseerd.

7 Discussie

De vraag of het teamgerichte nascholingsprogramma 'Omgaan met combinatieklassen', uitgevoerd binnen de reguliere setting van nascholing en begeleiding leidt tot verbeteringen in de management- en instructievaardigheden van leerkrachten kan bevestigend worden beantwoord.

Na afloop van de nascholing is de kwaliteit en kwantiteit van toepassing van beoogde onderwijsvaardigheden uit het directe instructiemodel verbeterd bij de leerkrachten die de cursus hebben gevolgd. Bovendien blijken ze hun instructie beter te organiseren en beter af te stemmen op het niveau van de leerlingen en op de eisen van de leertaak. Ook op de aspecten 'gebruik van materiaal en ruimte' en 'preventie en omgang met storingen' zijn verbeteringen opgetreden. Daarnaast blijkt dat deelname aan de teamgerichte nascholing heeft geleid tot een stijging in de taakgerichte leertijd bij de leerlingen.

Coaching heeft inderdaad voor twee aspecten van het onderwijsgedrag geleid tot extra effecten bovenop de nascholing: de organisatie van de instructie en het omgaan met storingen. Dit heeft mede geresulteerd in een sterkere groei in taakgerichtheid in de klassen van gecoachte leerkrachten.

Wanneer we deze effecten vergelijken met de effecten uit de eerste beproeving, dan blijkt dat voor de meeste onderscheiden criteriumvariabelen de effecten tijdens de eerste verbeteringsstudie groter waren dan die van deze studie (zie Veenman, Lem & Roelofs, 1990). De stijging in taakgerichtheid was in de eerste studie veel sterker, terwijl ook de verbetering in de management- en instructievaardigheden sterker was. Mogelijk biedt de grotere mate van 'ownership' en de grotere mate van bekendheid met de cursusinhoud tijdens de eerste beproeving een verklaring voor dit verschil. In de eerste verbeteringsstudie hadden de nascholers het programma zelf ontwikkeld en waren zij tevens onderzoeker. Mogelijk waren zij beter in staat de onderzoeksgegevens in de klas te verbinden aan de cursusinhoud, om zo de cursisten beter te richten op de te verbeteren vaardigheden. Uit studies naar effectiviteit van nascholing blijkt daarnaast dat programma's die geïnitieerd en geleid worden door een universitaire onderzoeker effectiever zijn dan programma's die door andere partijen zijn geïnitieerd (Lawrence & Harrison, 1980; Joslin, 1980; Wade, 1984). Niettemin blijkt het programma succesvol wanneer dit uitgezet wordt in de context van nascholing en begeleiding.

De resultaten van de observaties in klassen tonen aan dat gecoachte leerkrachten, hoewel klein in aantal, grotere vorderingen maakten op enkele eerder genoemde instructie- en managementvaardigheden dan niet-gecoachte leerkrachten. Mede hierdoor was de groei in taakgerichtheid groter in klassen van gecoachte leerkrachten dan in klassen van hun niet-gecoachte collega's. De resultaten vormen ten dele een bevestiging van de bevindingen van Bennett (1987) die in zijn meta-analyse een groot effect vond van coaching op de transfer van vaardigheden. Ook vormen de bevindingen een gedeeltelijke bevestiging van de resultaten van Showers (1982, 1984) en Baker (1983) in Amerikaanse experimenten. Echter, deze auteurs boekten resultaten die met de ons beschikbare gegevens niet getoetst konden worden. Zo vond men dat leerkrachten de aangeleerde vaardigheden niet alleen adequater toepasten, maar dat zij de vaardigheden ook langer vasthielden en ook een helderder beeld hadden van de verworven vaardigheden. Vervolgonderzoek zal uit dienen te wijzen of de gevonden effecten beklijven. De resultaten lijken overigens positiever dan die van Snippe (1991), die geen verschillen vond in implementatie tussen leerkrachten die wel en niet gecoacht waren.

De resultaten van deze studie laten zien dat verschillen in effecten van het totale programma (nascholing en coaching) terug te voeren zijn op verschillen in de wijze waarop de coaching is vormgegeven. Deze bevindingen sporen met die van Showers (1984). Zij vond eveneens dat de effectiviteit van de coaching afhankelijk was van een aantal vormgevingskenmerken: naarmate leerkrachten meer open stonden voor experimenteren en naarmate de coach meer aandacht besteedde aan experimenteren met nieuwe werkwijzen, werd transfer meer bevorderd.

Uit de resultaten kan worden opgemaakt dat de door Joyce en Showers (1988) genoemde functie van coaching 'analyse van toepassing' een onderscheidende rol heeft gespeeld in de effectiviteit van coaching. Deze functie betreft de mate waarin tijdens coaching aandacht wordt besteed aan de inpassing van een geleerde vaardigheid in de dagelijkse lespraktijk. Leerkrachten die sterk waren vooruitgegaan vonden dat hun coach veel praktische suggesties had gedaan, die bovendien toepasbaar bleken. Daarnaast vond men dat er sprake was geweest van een goede aansluiting bij de inhoud van de cursus.

De mate waarin feedback werd gegeven aan de hand van checklists bleek achteraf negatief samen te hangen met de vooruitgang in instructievaardigheden. Een verklaring hiervoor kan zijn, dat bij die leerkrachten, waar veelvuldig gebruik gemaakt is van dit soort lijsten, de coaching is blijven steken in een fase van vaststellen van omissies in onderwijsgedrag. Hierdoor kreeg de coaching mogelijk niet voldoende diepgang, gepaard gaande met een beperktere aandacht voor de inbedding van de nieuwe onderwijsgedragingen in het dagelijkse repertoir van de leerkracht. Joyce en Showers (1988) spreken in dit verband over fasen in het proces van coaching: een eerste fase van technische optimalisering van nieuwe vaardigheden en een tweede fase waarin de leerkracht de nieuwe vaardigheden definitief inpast in zijn lespraktijk.

Ten aanzien van de validiteit van uitspraken over de relatie tussen vormgeving en effecten van coaching gelden enkele beperkingen. Ten eerste zijn de bevindingen gezien de beperkte en mogelijk selectieve aard van de steekproef niet generaliseerbaar naar een grotere groep. Daarnaast is het mogelijk dat bij de vergelijking tussen de groepen leerkrachten die sterk en minder sterk vooruitgingen de eerste groep geneigd is geweest positiever te oordelen over de coaching dan de tweede groep. Tegen deze tweede beperking kan worden ingebracht dat zowel tijdstip als bron van de vaststelling van vormgevingskenmerken en effectiviteit verschilden. De vooruitgang in instructievaardigheden, uitgedrukt in verschilscores op de MIS, is vastgesteld door observatoren in de weken na afloop van de nascholing, terwijl leerkrachten zich direct na de nascholing hebben uitgesproken over de waarde van de coaching.

Met behulp van dit onderzoek is voorts aangetoond dat empirisch onderzoek naar succesvol onderwijsgedrag ten dienste gesteld kan worden van de professionalisering van leerkrachten en meer in het algemeen van teamgerichte nascholing. Dit resultaat mag verheugend worden genoemd, omdat de inhoud van veel nascholingsprogramma's nauwelijks kan bogen op een empirische basis (zie Guskey & Sparks, 1991).

Na de ervaringen opgedaan in deze studie en in de eerste beproeving van het programma OMC (Lem, Veenman, Nijssen & Roelofs, 1988) kan bovendien gesteld worden dat het door ons gehanteerde model van teamgerichte nascholing effectief is geweest. Het programma 'Omgaan met combinatieklassen' kan beschouwd worden als een integratie van een aantal effectief gebleken modellen voor teamgerichte nascholing. Door middel van training kan de beheersing van nieuwe onderwijsvaardigheden worden bevorderd. Behalve gerichte training bevat het model voor teamgerichte nascholing tevens een klinisch-collaboratieve component (Sparks, 1983a,b): leerkrachten onderzochten samen met de coach het proces van leren en onderwijzen. Uitgangspunt hierbij was dat de leerkrachten reeds beschikken over een grote hoeveelheid kennis en dat groei in onderwijsvaardigheden plaatsvindt door middel van het bestuderen van en reflecteren op het eigen lesgedrag.

Tot slot dient een aantal kanttekeningen te worden gemaakt bij dit onderzoek. Gecoachte en niet-gecoachte leerkrachten bleken op enkele relevante aspecten significant van elkaar te verschillen op de voormeting. Dit kan de in-

terpretatie van de effecten bemoeilijken. Mogelijk is toch sprake geweest van een zelf selectie-effect. Bij navraag bleek dat voor de conditie 'nascholing met coaching' niet bewust gekozen was voor leerkrachten met specifieke problemen. Een vraag die nog beantwoord moet worden is of en in welke mate de door de leerkrachten verworven vaardigheden zullen beklijven. Dit is een indicator van institutionalisering. In dit onderzoek stond de fase van implementatie en transfer centraal. Over de daarop volgende fase - de fase van de institutionalisering waarin sprake is van bestendiging van gevonden effecten - kan op dit moment geen uitspraak worden gedaan. Momenteel wordt een vervolgonderzoek uitgevoerd naar de mate waarin de effecten van het nascholingsprogramma standhouden.

Literatuur

ARBO (1984). *Nieuw, concreet en zichtbaar.* Zeist: Onderwijscentrum.

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.

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

Fullan, M.G. (1990). Staff development, innovation and institutional development. In B. Joyce (Ed.), *Changing school culture through staff development* (pp. 3-25). Alexandria, VA: Association for Supervision and Curriculum Development.

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

Guskey, T.R., & Sparks, D. (1991). What to consider when evaluating staff development. *Educational Leadership, 48,* 73-76.

Joslin, P.A. (1980). *Inservice teacher education: A meta-analysis of the research.* Unpublished doctoral dissertation, University of Minnesota.

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

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

Lawrence, G., & Harrisson, D. (1980). Policy implications of the research on the professional development of education personnel: An analysis of fifty-nine studies. In C.E. Freistritzer (Ed.), *The 1981 Report on Educational Personnel Development* (pp. 145-168). Washington, DC.: Feistritzer Publications.

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

Roelofs, E.C. (1993). *Teamgerichte nascholing en coaching: een experimentele studie in scholen met combinatieklassen.* Academisch proefschrift Katholieke Universiteit Nijmegen.

Roelofs, E., Gennip, H. van, & Veenman, S. (1989). *Zoeklicht op combinatieklassen met kleuters.* Nijmegen: Facultaire Vakgroep Interdisciplinaire Onderwijskunde.

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

Showers, B. (1982). *Transfer of training: the contribution of coaching.* Eugene, Oregon: Center for Educational Policy and Management.

Showers, B. (1984). *Peer coaching: a strategy for facilitating transfer of training.* Eugene, OR: Center for Educational Policy and Management, College of Education, University of Oregon.

Simons, P.R.J. (1990). *Transfervermogen.* Inaugurele rede. Nijmegen: Katholieke Universiteit Nijmegen.

Snippe, J. (1988). Effecten van klasseconsultatie op curriculumimplementatie van leerresultaten. *Tijdschrift voor Onderwijsresearch, 13,* 141-150.

Snippe, J. (1991). *In-service training voor leerkrachten.* Academisch proefschrift Rijksuniversiteit Groningen.

Sparks, D., & Loucks-Horseley S. (1990). Models of staff development. In W.R. Houston (Ed.), *Handbook of research on teacher education* (pp. 234-250). New York: Macmillan.

Sparks, G.M. (1983a). *Inservice education: training activities, teacher attitude, and behavior change.* Unpublished doctoral dissertation, Stanford University.

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

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

Tulder, M. van, Veenman, S., & Roelofs, E. (1989). *Nascholing leraren basisonderwijs. Vormgeving en resultaten.* 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs (Selecta Reeks).

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

Veenman, S., Lem, P., & Nijssen, F. (1988, 1990). *Omgaan met combinatieklassen: Een programma voor schoolverbetering.* s' Gravenhage: Instituut voor Onderzoek van het Onderwijs (Selecta Reeks).

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. (1993). *Effectieve instructie en doelmatig klassemanagement: Een schoolverbeteringsprogramma voor enkelvoudige en combinatieklassen.* (4de druk) Amsterdam/Lisse: Swets & Zeitlinger.

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

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

Manuscript aanvaard 7-7-1993

Auteurs

E. Roelofs is universitair docent bij de Vakgroep Onderwijskunde van de Universiteit Utrecht.

S. Veenman is universitair hoofddocent bij de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

J. Raemaekers is wetenschappelijk medewerker bij de Vakgroep Onderwijskunde van de Katholieke Universiteit Nijmegen.

Correspondentie-adres: Universiteit Utrecht, Vakgroep Onderwijskunde, Postbus 80140, 3508 TC Utrecht.

Abstract

Staff development and coaching: Effects on teaching behaviour in mixed-age classes

E. Roelofs, S. Veenman & J. Raemaekers. Pedagogische Studiën, 1993, 70, 352-367.

Based on the results of observations in mixed-age classes a staff development programme 'Dealing with Mixed-age Classes' (DMC) was developed by the University of Nijmegen. In 1990 the staff development programme was implemented by teacher educators and school counsellors in the context of the regular educational support structure.

DMC was implemented at five sites in The Netherlands by teacher educators and school counsellors. A total of 89 teachers from 12 schools participated in the in-service training.

A quasi-experimental design was used to test the effects of DMC and coaching, using two experimental groups (coached n=18; uncoached n=10) and one control group (n=14). Based on pre- and posttraining classroom observations, a significant treatment effect was found for pupil's time-on-task levels in mixed-age classrooms, and for teacher instructional and classroom management skills. Coached teachers improved more strongly on two aspects of classroom organization: organizing effective instruction and dealing with disturbances. Finally, effects of coaching were related to the way coaching had actually been performed.

Het belang van informatie over een afwezige genetische vader*

H. M. Y. Koomen, J. B. Hoeksma en M. Meerum Terwogt**

Samenvatting

In een gestructureerd interview werden 102 volwassen proefpersonen, allen opgegroeid met een vervangende vader, ondervraagd over de informatie die ze over hun afwezige genetische vader hadden. Daarnaast werden, zowel met mondelinge vragen als met twee schriftelijke vragenlijsten (NPV en Erfelijkheidsvragenlijst), diverse aspecten van het huidige persoonlijk functioneren van deze personen onderzocht. De belangrijkste bevindingen waren: 1. de hoeveelheid en kleur van de informatie waarover personen zeggen te beschikken speelt een minder cruciale rol in het persoonlijk functioneren dan verwacht; 2. een ontmoeting met de genetische vader blijkt zowel een gunstig effect te hebben op geïnformeerdheid als op het latere sociale functioneren; 3. het blijkt essentieel onderscheid te maken naar de reden van afwezigheid van de genetische vader: kinderen geboren uit een kortstondige relatie zijn zowel wat betreft geïnformeerdheid als persoonlijk functioneren slechter af dan kinderen van overleden vaders en gescheiden ouders, waarbij de laatste groep een middenpositie inneemt.

1 Inleiding

De vraag naar het belang van informatie over een ontbrekende genetische vader is niet zonder betekenis. Zij speelt bijvoorbeeld een belangrijke rol in het publieke debat over het opheffen van donoranonimiteit bij kunstmatige inseminatie (zie bijvoorbeeld Bierkens, 1992; Meerum Terwogt, 1991; 1992). En daarnaast zien we haar ook terug in beschouwingen over adoptie, waarin de aandacht uitgaat naar het verschijnsel 'afstammingsonrust' (Hoksbergen, 1990). Dit artikel doet verslag van een onderzoek naar de geïnformeerdheid over afwezige genetische ouders. Aandacht wordt besteed aan een tweetal vragen: welke variabelen zijn van invloed op de *hoeveelheid, aard* en *kleur* van de informatie die het kind over een afwezige genetische ouder krijgt toegespeeld? En, hoe werkt deze informatie door bij het betreffende individu? Het verslag betreft een retrospectieve interviewstudie bij personen, bij wie het contact tussen de genetische ouders om de één of andere reden zo vroeg verbroken is, dat zij de genetische vader niet hebben leren kennen. Om uit te kunnen sluiten dat allerlei verschijnselen mede zouden zijn terug te voeren op de afwijkende gezinssituatie die het ontbreken van een vaderfiguur met zich meebrengt, zijn uitsluitend die personen in het onderzoek betrokken die van jongs af aan in een stabiel gezin met een vervangende vader zijn opgegroeid. Onderzoek in stiefgezinnen toont immers aan dat zo'n sociale vader in veel opzichten, zij het niet in alle (Amato & Keith, 1991), als een volwaardige vervanger kan worden beschouwd (Amato, 1987; Oshman, 1975).

Voor wij de beide hoofdvragen toelichten, is een algemene opmerking op zijn plaats: als we het hebben over de hoeveelheid, aard of kleur van de informatie, bedoelen we steeds de *subjectieve beleving* van die aspecten. De feitelijke hoeveelheid en kleur van de informatie is met behulp van een retrospectief interview nauwelijks betrouwbaar te bepalen. Zo spelen soms fantasieën van de ondervraagde een niet geringe rol. Hoewel de beleving ongetwijfeld niet losstaat van de feitelijke gang van zaken,

* Dit onderzoek werd mede mogelijk gemaakt door financiële ondersteuning van het Nationaal Fonds voor de Geestelijke Volksgezondheid, de Stichting Kinderpostzegels Nederland en het Programmeringscollege Onderzoek Jeugd (WVC).

** Wij bedanken C. Reijnders, L. van Kampen en E. van Kasteel voor hun bijdrage aan de dataverzameling.

kan ze onder invloed van dergelijke factoren aanzienlijk afwijken. Echt bezwaarlijk kan het ontbreken van feitelijke informatie overigens niet genoemd worden. Bij de vraag naar het psychisch welbevinden van mensen is de subjectieve beleving van de situatie immers van doorslaggevend belang.

Factoren van invloed op geïnformeerdheid

Wanneer naar achtergronden van de informatie of kennis over de genetische vader wordt gezocht, ligt het in de eerste plaats voor de hand aandacht te besteden aan de *oorzaak of reden van afwezigheid* van de genetische vader. Op diverse plaatsen in de literatuur worden nakomelingen die een ouder missen door *overlijden* vergeleken met kinderen van *gescheiden* ouders. Uit een meta-analyse van Amato en Keith (1991) komt naar voren dat kinderen waarvan één van beide ouders is overleden weliswaar slechter af zijn dan kinderen uit reguliere twee-oudergezinnen, maar dat overlijden toch een minder negatief effect heeft dan afwezigheid door echtscheiding. Zo blijkt in het laatste geval, de attitude ten opzichte van de vader beduidend negatiever dan bij overlijden (Hetherington, 1972). Dit wordt begrijpelijk als we bedenken dat echtscheiding door kinderen als een vrijwillige daad gezien wordt, waardoor het meer dan overlijden als een verlies van liefde wordt ervaren (Tusenius, 1989). Daarnaast is aannemelijk dat ook de door moeder verstrekte informatie haar werk doet. Hainline and Feig (1978) vonden in dit verband dat gescheiden ouders negatiever over elkaar spraken dan weduwen over hun overleden man. De laatsten waren meer vergelijkbaar met niet-gescheiden ouders. Behalve overlijden en echtscheiding, in welke gevallen de genetische vader toch ooit in zekere zin een ouderrol heeft vervuld, kan de afwezigheid van de genetische vader ook een heel andere achtergrond hebben. Zo bestaat er een aanzienlijke groep nakomelingen voor wie de genetische vader eigenlijk nooit een rol van betekenis heeft gespeeld. Het betreft hier gevallen waarin een kind werd verwekt in een *kortstondige buitenechtelijke c.q. voorechtelijke relatie* of een eenmalig contact. Naar deze laatste categorie als aparte groep is, naar wij weten, nauwelijks onderzoek gedaan. Het is aannemelijk dat de situatie voor deze kinderen nog weer heel anders ligt.

Een tweede variabele waarvan zich een effect zou kunnen voordoen is het *tijdstip* waarop de persoon over het bestaan van zijn of haar genetische vader is geïnformeerd. In een exploratief onderzoek bij geadopteerden die op zoek gingen naar informatie over hun genetische ouders, bleek dat degenen die voor hun 10de jaar geïnformeerd waren over hun adoptiestatus zich in een aantal opzichten positief onderscheidden van personen die later op de hoogte waren gesteld (Triseliotis, 1973). Later geïnformeerden gaven niet alleen aan veel meer geschokt en verward te zijn geweest door het bericht, ze vertelden ook wantrouwen te koesteren ten aanzien van de verstrekte informatie (Triseliotis, 1973). Personen die later geïnformeerd zijn zouden zich om deze redenen met dezelfde informatie minder goed geïnformeerd kunnen voelen dan vroeg geïnformeerden.

Een derde variabele, die bijna vanzelfsprekend een rol speelt, is het wel of niet hebben *ontmoet* van de genetische vader. Van de geadopteerden die op latere leeftijd op zoek gaan naar hun genetische ouders heeft een groot deel volgens deskundigen als motief iets essentieels toe te voegen aan de informatie die men tot dan toe heeft (Hoksbergen, 1992). In het exploratieve onderzoek van Triseliotis (1973) was de meerderheid van de geadopteerden inderdaad van mening dat na het gerealiseerde contact een nieuw stadium was bereikt. Hoewel de ontmoeting met een genetische ouder zelden voldeed aan de vooraf gekoesterde verwachtingen, vonden de meesten achteraf toch dat deze confrontatie met de realiteit hen goed had gedaan: ze hadden een bepaalde rust gevonden. Amato en Keith (1991) onderzochten in hun meta-analyse of kinderen van gescheiden ouders baat hebben bij contact met de afwezige ouder. Zij vonden echter maar beperkte ondersteuning voor hun hypothese. In een aantal studies werd een gunstige invloed van contact vastgesteld, in andere studies geen invloed en in weer andere leidde veel contact juist tot meer problemen. Niettegenstaande deze resultaten, zou in de lijn van de bevindingen van Triseliotis verwacht mogen worden dat een ontmoeting met de genetische vader bijdraagt aan het gevoel goed geïnformeerd te zijn.

Effecten van geïnformeerdheid

Ook ten aanzien van de tweede vraag is het zinvol ons licht op te steken bij de adoptieliteratuur. In deze literatuur komt naar voren dat een gebrek aan openhartigheid omtrent het genetische ouderschap een nadelige invloed heeft op de persoonlijkheids- en identiteitsontwikkeling (Hoksbergen, 1987; Raad voor het Jeugdbeleid, 1988; Sibbing-Willems & Temmink-Spits, 1991). Daarbij ziet men zich gesteund door een aantal empirische bevindingen: adoptiekinderen blijken een verhoogde kans te hebben op het ontwikkelen van psycho-pathologie (Verhulst & Versluis-den Bieman, 1989) en worden relatief vaak aangemeld bij instellingen voor jeugdhulpverlening (Hoksbergen & Walenkamp, 1983). De persoonlijkheids- en gezinsproblematiek van deze kinderen blijkt bovendien gerelateerd aan het zoeken naar informatie over de genetische ouders (Triseliotis, 1973; Goebel & Lott, 1986; Stein & Hoopes, 1985), hoewel hierbij aangetekend moet worden dat niet duidelijk is hoe de relatie tussen psycho-pathologie en zoekgedrag in elkaar zit (Slijper & Verhulst, 1991). Zo is bijvoorbeeld onbekend in hoeverre gedragsproblemen door onbekende vroege perinatale- en ervaringsfactoren hierbij een rol spelen. Concrete beschrijvingen van probleemgedrag die we in de adoptieliteratuur tegenkomen zijn onder andere: negatief zelfbeeld, geremdheid in contacten met anderen, gevoelens van depressie, gespannenheid en angst (Triseliotis, 1973), afwijkende identiteitsscores (Stein & Hoopes, 1985), afwijkende persoonlijkheidsscores (Goebel & Lott, 1986) en slechte gezinsrelaties (allen). Ook bij kinderen uit een-oudergezinnen met een scheidings- dan wel overlijdenshistorie, en bij kinderen uit stiefgezinnen vinden we een aantal van deze probleemgedragingen terug. Amato en Keith (1991) vonden voor deze groepen in vergelijking met kinderen uit reguliere gezinnen significant meer problemen op het gebied van gedrag (agressie, delinquentie e.d.), emotioneel functioneren (depressie, angsten e.d.), zelfbeeld en sociale relaties (niet bij overlijden). Wat betreft emotioneel functioneren waren kinderen uit stiefgezinnen zelfs slechter af dan kinderen uit gescheiden een-oudergezinnen. Op grond van deze bevindingen bij diverse typen gezinnen lijkt het van belang ook bij onze doelgroep dergelijke met persoonlijk functioneren samenhangende variabelen te onderzoeken. In dit verband is het onder meer interessant na te gaan of geïnformeerdheid implicaties heeft voor diverse persoonlijkheidskenmerken, voor tevredenheidsgevoelens ten aanzien van diverse levensgebieden (bijvoorbeeld sociaal functioneren) en voor eventuele psychische problemen.

Daarnaast is het van belang te onderzoeken of de informatie invloed heeft op hoe genetische vaders hun nakomelingen in het latere leven nog bezighouden. Heeft geïnformeerdheid effect op *gedachten over de genetische vader* in het latere leven? Kinderen die gescheiden leven van hun vader blijken de neiging te hebben in hun fantasie een ideaalbeeld van de vader te construeren (Bach, 1946). Irreële fantasieën, zowel in idealistische als negatieve zin, vormen ook een bekend verschijnsel in de klinische adoptiepraktijk (Sibbing-Willems & Temmink-Spits, 1991). In een aantal gevallen is het een hardnekkig fenomeen dat tot ver in de volwassenheid blijft bestaan. In het exploratieve onderzoek van Triseliotis (1973) bleven vooral geadopteerden die slecht of negatief geïnformeerd waren over hun achtergrond sterk gepreoccupeerd met hun genetische ouders. In dit licht zou ook bij dit onderzoek een effect van geïnformeerdheid op gedachten en fantasieën over de afwezige genetische vader verwacht kunnen worden.

Mensen die hun genetische ouders niet beide kennen blijken meer belang te hechten aan erfelijke factoren dan de gemiddelde Nederlander geneigd is te doen (Meerum Terwogt, Hoeksma & Koops, 1991). In aansluiting bij de gedachtengang van Van Kampen, Koops, Meerum Terwogt en Reijnders (1990) is daarom tevens nagegaan of de geïnformeerdheid over de genetische vader van invloed is op de *erfelijkheidsopvatting* van het individu. Volgens genoemde auteurs kunnen de irreële, positief getinte fantasieën over de ontbrekende vader begrepen worden als het produkt van een attributieproces dat optreedt ten tijde van de identiteitsvorming. In dit proces worden individuele gedragingen en gedachten, die niet herleidbaar zijn tot de moeder en een eventueel plaatsvervangende vader, toegeschreven aan de onbekende genetische vader, waarmee erfelijke factoren een grote betekenis wordt toegekend. In een kleinschalig vragenlijstonderzoek

vonden Van Kampen et al. (1990) empirische ondersteuning voor deze idee: personen die slecht en laat geïnformeerd waren, bleken sterker te geloven in de erfelijkheid van psychologische eigenschappen. Ook Triseliotis (1973) nam in zijn onderzoek bij geadopteerden, een sterke erfelijkheidsopvatting waar. Anders dan voornoemde auteurs legt hij ook expliciet de relatie met *negatieve* informatie over de genetische ouders: "Most of them being firm believers of heredity, they believed that the 'badness' of their parents might be transmitted to them and through them to their children" (1973, p.49). In navolging van genoemde auteurs verwachten wij dat personen die zich niet goed geïnformeerd voelen een naar verhouding sterke erfelijkheidsopvatting hebben.

2 Methode

2.1 Proefpersonen en procedure

Aan het onderzoek namen 102 personen deel. De proefpersonen werden geworven via een annonce onder de categorie 'oproepen' in de grote landelijke dagbladen (i.c. De Volkskrant, Trouw, Het Parool en De Telegraaf). De oproep was gericht aan mensen die tijdens hun jeugd hun biologische vader niet hadden gekend en die bereid waren ten behoeve van een wetenschappelijk onderzoek anoniem hun ervaringen te vertellen. De volgende drie criteria werden hierbij vermeld: a) de respondent diende zelf 21 jaar of ouder te zijn; b) de genetische vader moest voor het 5e jaar uit het leven van de respondent verdwenen zijn en indien zich later een ontmoeting had voorgedaan mocht dit niet voor het 21e jaar gebeurd zijn; en c) de respondent diende opgevoed te zijn door de genetische moeder en een vervangende vader. Bij de (telefonische) aanmelding werden deze criteria gecheckt. Er werd tevens nagegaan of er sprake was geweest van een stabiele en normale opvoedingssituatie waarin de vaderrol volwaardig was vervuld. Indien bleek dat de sociale vader pas geruime tijd na het wegvallen van de genetische vader zijn intrede had gedaan, dan wel dat er sprake was geweest van wisselende 'vaders', werd de aanmelder niet in het onderzoek betrokken. Hetzelfde gebeurde met respondenten met traumatische ervaringen (vermoedens van incest, mishandeling, verwaarlozing etc.). Op grond van de laatste twee criteria moest 10% van de aanmelders van deelname worden uitgesloten. Van de deelnemers bleek het merendeel van het vrouwelijk geslacht (66%). De gemiddelde leeftijd was 36 jaar (*sd*=8.5). De helft (52%) was hoogopgeleid (universitair of hbo-niveau), 36% had een middelbare en 12% een lagere opleiding. Ruim de helft was gehuwd of woonde samen (55%), 35% noemde zichzelf alleenwonend en 10% was gescheiden. De oververtegenwoordiging van hoger opgeleiden moet vermoedelijk worden toegeschreven aan het feit dat mensen met een hogere opleiding doorgaans meer bereid zijn aan wetenschappelijk onderzoek mee te werken. Het naar verhouding hoge percentage vrouwen hangt waarschijnlijk samen met de grotere bereidheid en gelegenheid die vrouwen doorgaans hebben om over persoonlijke aangelegenheden te praten. Bij iedere deelnemer werd een gestructureerd interview (Van Kampen et al., 1990) en een tweetal schriftelijke vragenlijsten afgenomen.

2.2 Materiaal

Het interview bevatte ruim 100 vragen, waarvan wij er voor dit onderzoek 31 gebruikten. Zeven vragen hadden betrekking op de hoeveelheid en drie op de kleur van de informatie over de genetische vader. Er werden vier vragen gesteld over de mate van tevredenheid op het gebied van werk, sociale contacten, het huidige leven en de relatie met de andere sekse. In twee vragen werd informatie ingewonnen over eventuele hulp gezocht voor psychische problemen. De overige vragen betroffen de reden van afwezigheid van de genetische vader, het tijdstip van informeren over het bestaan van de genetische vader, het al dan niet hebben ontmoet van de genetische vader, fantaseren over de genetische vader en ideeën over gelijkenis met de vader qua uiterlijk en karakter. Ten slotte bevatte het interview een aantal vragen over de hierboven vermelde demografische gegevens. De twee vragenlijsten die werden afgenomen waren de NPV (Luteijn, Starren & Van Dijk, 1975) en de Erfelijkheidsvragenlijst (Meerum Terwogt et al., 1991). Beide lijsten werden door de proefpersonen zelf ingevuld. De NPV, bestaande uit 133 items verdeeld over zeven schalen, wordt verondersteld de volgende psychologische kenmerken te meten:

inadequatie, sociale inadequatie, rigiditeit, verongelijktheid, zelfgenoegzaamheid, dominantie en zelfwaardering. De Erfelijkheidsvragenlijst, bestaande uit 19 items, beoogt te meten in hoeverre een persoon algemeen menselijke kenmerken van uiteenlopende aard (netheid, depressiviteit, hoogtevrees, creativiteit etc.), aan erfelijke dan wel omgevingsfactoren toeschrijft. Per kenmerk dient dit op een vijfpuntsschaal te worden aangegeven. Zowel de NPV (Luteijn et al., 1975) als de Erfelijkheidsvragenlijst (Meerum Terwogt et al., 1991) zijn in eerder onderzoek betrouwbare instrumenten gebleken.

3 Resultaten

3.1 Hoeveelheid en kleur van de informatie

De vragen die betrekking hadden op het *kwantitatieve* aspect van de informatie-overdracht en de beschikbare kennis over de genetische vader, waren de volgende: hoeveel kennis had u als jongere (13-20 jaar) over uw genetische vader? Hoeveel kennis heeft u nu? Had u alleen feitelijke kennis over uw vader of had u ook kennis over zijn persoonlijkheid? Hoe is dat op dit moment? Kwam u makkelijk of moeilijk aan die kennis? Verstrekte uw moeder deze informatie gemakkelijk of was ze in dit opzicht terughoudend? Hoe is dat nu? Tabel 1 geeft een overzicht van de verdeling van de antwoorden over de diverse antwoordcategorieën. De aantallen overziend blijkt dat het merendeel van de ondervraagden, zowel tijdens de adolescentie als op het moment van ondervraging, volgens eigen zeggen over weinig of geen informatie over de genetische vader beschikte. Voor ongeveer tweederde deel van de geïnterviewden geldt bovendien dat hun moeder in dit opzicht terughoudend was of dat ze zelfs in het geheel geen informatie over de genetische vader verstrekte. Desalniettemin zegt iets meer dan de helft van de ondervraagden momenteel, behalve over feitelijke kennis, ook over informatie over de persoonlijkheid van de vader te beschikken. Dit laatste wordt begrijpelijk als we erbij betrekken dat ruim de helft van de ondervraagden naast de moeder nog andere informatiebronnen had en dat eenderde deel op latere leeftijd zelf contact heeft gezocht met de genetische vader.

De antwoorden op de vragen, zoals in Tabel 1 weergegeven, werden aan een Homals-analyse onderworpen. Doel van de analyse was te ko-

Tabel 1
Resultaten Homals-analyse op vragen betreffende de hoeveelheid informatie

Vraag	Antwoord	Freq.	Schaalwaarde	Discr.waard
Hoeveel kennis tijdens adolescentie	veel	27	1.14	.66
	weinig	51	-.41	
	geen	24	-1.51	
Hoeveel kennis nu	veel	41	.68	.39
	weinig	53	-.41	
	geen	5	-1.51	
Aard v.d. kennis tijdens adolescentie	persoon	42	.90	.51
	feitelijk	51	-.59	
Aard v.d. kennis nu	persoon	59	.55	.44
	feitelijk	39	-.83	
Hoe kwam u aan de kennis	makkelijk	38	1.01	.63
	moeilijk	64	-.63	
Informatieverstrekking moeder toen	makkelijk	32	1.03	.56
	moeilijk	45	-.27	
	niet	25	-.89	
Informatieverstrekking moeder nu	makkelijk	35	.94	.49
	moeilijk	41	-.47	
	niet	23	-.65	

men tot een schaal, die de gemeenschappelijke aspecten in de antwoorden op de verschillende vragen representeert. Hiertoe werd één dimensie geëxtraheerd. De eigenwaarde, die in dit geval als de fractie verklaarde variantie kan worden opgevat, bedroeg 0.52. Kolom vier bevat de met Homals berekende categorie-kwantificaties (de aan de antwoordcategorieën toegekende schaalwaarden). Kolom vijf bevat de discriminatie-waarden die de correlaties tussen de vragen en de gevonden dimensie weerspiegelen (vergelijkbaar met factorladingen). De categoriekwantificaties zijn goed interpreteerbaar. De kwantificaties van de antwoorden die op veel informatie en informatieverstrekking zonder terughouding duiden, alsmede op informatie over de persoonlijkheid van de genetische vader zijn alle positief, terwijl de overige antwoorden die op het tegendeel duiden negatief zijn. Op grond van de gepresenteerde schaalwaarden lijkt het dan ook gerechtvaardigd om de resulterende schaal op te vatten als een maat voor de hoeveelheid informatie over de genetische vader, waarover een persoon zegt te beschikken.

Het interview bevatte drie vragen over *kwalitatieve aspecten* van de informatieoverdracht en de beschikbare kennis, namelijk: wat voor beeld had u van uw genetische vader in de periode van 13 tot 20 jaar? Hoe denkt uw moeder over uw genetische vader? Hoe bracht uw moeder de informatie over uw genetische vader? In alle drie vragen gaat het om de *kleur* van percepties of informatie over de genetische vader met antwoordalternatieven in termen van positief en negatief. Tabel 2 bevat de frequenties waarmee diverse antwoorden werden gegeven. Ongeveer 60% van de ondervraagden geeft aan tijdens de adolescentie voornamelijk positief over de genetische vader te hebben gedacht. Naar het oordeel van de ondervraagden heeft eveneens een kleine 60% van de moeders een positief beeld van de vader. Volgens ongeveer 30% van de ondervraagden verstrekte de moeder voornamelijk positieve informatie over de genetische vader.

Om tot een schaal voor de kleur van de informatie te komen werd op de antwoorden van de drie genoemde vragen een Homals-analyse uitgevoerd. De analyse resulteerde in een eendimensionale oplossing met een eigenwaarde van 0.71. De categorie-kwantificaties alsmede de discriminatie-maten behorende bij de drie vragen staan weergegeven in Tabel 2. Alle positieve antwoord-categorieën blijken positieve kwantificaties (schaalwaarden) te hebben. De overige categorieën zijn minder consistent geordend. De resulterende schaal kan dan ook het beste worden opgevat als een maat voor de kleur van de informatie, waarbij hoge scores als positief worden opgevat en lage scores als niet positief.

3.2 Factoren die de hoeveelheid en kleur van de informatie bepalen

De afwezigheid van de genetische vader heeft niet bij alle geïnterviewden eenzelfde oorzaak.

Tabel 2
Resultaten Homals-analyse op vragen betreffende de kleur van de informatie

Vraag	Antwoord	Freq.	Schaalwaarde	Discr.-waarde
Beeld vader tijdens adolescentie	sterk positief	32	.66	.53
	positief	19	.76	
	gemengd	20	-1.08	
	negatief	5	-.10	
	sterk negatief	10	-.68	
Hoe dacht moeder over vader	sterk positief	27	1.21	.86
	positief	23	.15	
	gemengd	6	-1.28	
	negatief	12	-1.42	
	sterk negatief	20	-.79	
Hoe bracht moeder de informatie	sterk positief	11	1.64	.73
	positief	16	1.00	
	objektief	25	.42	
	negatief	14	-1.14	
	sterk negatief	19	-.50	

In 15 van de 102 gevallen is er sprake van overlijden van de genetische vader, in 32 gevallen van echtscheiding en in 55 gevallen is de geïnterviewde geboren uit een kortstondige relatie dat wil zeggen uit een kortdurende voorechtelijke c.q. buitenechtelijke verhouding of een eenmalig contact. Door middel van twee *eenwegs* variantie-analyses werd nagegaan of de hoeveelheid en kleur van de informatie samenhangt met de reden van afwezigheid van de genetische vader. Tabel 3 laat de gemiddelden zien van de met Homals bepaalde scores voor de hoeveelheid en kleur van informatie in de drie onderscheiden groepen *overlijden, echtscheiding* en *kortdurende relatie*. Zowel de hoeveelheid als de kleur van de informatie vertoont een significante samenhang met de reden van afwezigheid van de genetische vader (hoeveelheid: $F(2,99)=5.56$, $p<0.001$; kleur: $F(2,97)=3.27$, $p<0.05$). Personen van wie de vader is overleden, blijken zich beter geïnformeerd te voelen dan personen van wie de vader ontbreekt door echtscheiding. De laatsten beschikken op hun beurt weer over meer informatie dan personen die geboren zijn uit een kortstondige relatie. De kleur van de informatie blijkt het positiefst in geval van overlijden en het minst positief bij echtscheiding.

Tabel 3

Hoeveelheid en kleur van de informatie in samenhang met de reden van afwezigheid van de vader: gemiddelden en standaarddeviaties

Reden		Hoeveelheid	Kleur
overlijden	*n*=15	0.60 (0.99)	0.49 (0.97)
echtscheiding	*n*=32	0.19 (0.96)	-0.36 (1.0)
kortdurende relatie	*n*=55	-0.29 (0.99)	0.02 (1.09)

De leeftijd waarop de geïnterviewden over het bestaan van hun genetische vader werden geïnformeerd varieert sterk tussen personen. Een flink aantal (36%) wist het vóór het vijfde jaar, 22% hoorde het tussen zes en twaalf jaar, 19% tussen dertien en zestien, 12% tussen zeventien en vierentwintig en 12% pas na vierentwintig jaar. De relatie tussen de hoeveelheid en kleur van de informatie enerzijds en de leeftijd waarop de ondervraagden werden geïnformeerd anderzijds werd geanalyseerd met regressie-analyse. Daarbij is rekening gehouden met mogelijke kromlijnige verbanden. Voor de hoeveelheid informatie werd een kromlijnig verband met informatie-leeftijd gevonden ($F(99,2)=5.91$, $p<0.005$, $r=0.33$). Nadere bestudering van de regressievergelijking leert dat de hoeveelheid informatie waarover personen zeggen te beschikken afneemt naarmate men later geïnformeerd is, om na een informatie-leeftijd van ongeveer zeventien jaar weer licht te stijgen. Met betrekking tot de kleur van de informatie werd een dergelijke significante relatie met leeftijd van informeren niet gevonden.

Zevenendertig ondervraagden gaven aan één of meer ontmoetingen te hebben gehad met hun genetische vader. Het blijkt dat zij die hun genetische vader daadwerkelijk hebben ontmoet, zich beter geïnformeerd voelen dan personen die hem niet hebben ontmoet ($F(1,97)=12.22$, $p<0.001$). Ontmoeting met de genetische vader had geen effect op de kleur van de informatie ($F(1,95)=0.25$, n.s.).

Ten slotte werd nagegaan of er een verband bestaat tussen de feitelijke leeftijd alsmede het geslacht van de geïnterviewden enerzijds en de hoeveelheid en de kleur van de informatie anderzijds. Noch voor geslacht, noch voor leeftijd werden significante relaties vastgesteld.

Tot zover werd de invloed van de variabelen *reden van afwezigheid, leeftijd van informeren, ontmoeting met de genetische vader, leeftijd* en *sekse* ieder afzonderlijk geanalyseerd. Om de gemeenschappelijke invloed van deze variabelen op de hoeveelheid en kleur van de informatie na te gaan voerden we een tweetal stapsgewijze regressie-analyses uit waarin alle genoemde factoren werden betrokken. De analyse met *hoeveelheid* informatie als responsvariabele resulteerde in drie significante predictoren ($r=0.53$), te weten: de variabele die aangeeft of er een ontmoeting met de genetische vader heeft plaatsgevonden ($b=0.91$, $s.e.(b)=0.19$, $(\beta=-0.43)$, $p<0.001$), de variabele die aangeeft of de vader ontbreekt door overlijden ($b=0.95$, $s.e.(b)=0.28$, $(\beta=0.32)$, $p<0.01$) en de leeftijd waarop de geïnterviewde is geïnformeerd ($b=-0.15$, $s.e.(b)=0.07$, $(\beta=-0.21)$, $p<0.05$). De geschatte regressiegewichten duiden erop dat het gevoel goed geïnformeerd te zijn groter is als er een feitelijke ontmoeting met de vader heeft plaatsgevonden of als de vader ontbreekt door overlijden en dat dit gevoel kleiner wordt naarmate de persoon op latere leeftijd is geïnformeerd. Betreffende de *kleur*

van de informatie resulteerde de analyse in één significante predictor, namelijk de variabele die aangeeft dat de vader ontbreekt door echtscheiding (b=-.49, $s.e.(b)$=0.23, (β=-0.21), p<0.05). Zoals we al eerder vaststelden is de kleur van de informatie over de genetische vader waarover de ondervraagden zeggen te beschikken in dit geval minder positief.

3.3 Effecten van hoeveelheid en kleur van de informatie

In het onderzoek naar effecten van hoeveelheid en kleur van de informatie worden drie groepen variabelen betrokken: a) variabelen die samenhangen met persoonlijk functioneren, te weten persoonlijkheidskenmerken zoals geoperationaliseerd met de NPV, tevredenheid op diverse levensgebieden en een variabele die aangeeft of er al dan niet hulp gezocht is voor psychische problemen; b) variabelen die verwijzen naar gedachten over de genetische vader in het huidige leven, namelijk fantaseren over de genetische vader en opvattingen over gelijkenis met de genetische vader qua uiterlijk en karakter; en c) erfelijkheidsopvattingen zoals geoperationaliseerd met de Erfelijkheidsvragenlijst.

In de vorige paragraaf werd van een aantal variabelen de invloed op hoeveelheid en kleur van de informatie onderzocht. Vanwege een mogelijke directe of indirecte invloed op de hierboven genoemde effectvariabelen, zijn ook deze achtergrondvariabelen in de analyses betrokken.

Persoonlijkheidskenmerken

De NPV bevat een zevental schalen die betrekking hebben op evenzovele persoonlijkheidskenmerken. Voor ieder van deze schalen berekenden we correlaties met de dimensies hoeveelheid en kleur van de informatie. Slechts één correlatie bleek significant, doch zwak, te weten die tussen de *kleur* van de informatie en de *rigiditeit-schaal* (r=-0.20, p<0.05). Naarmate personen zich negatiever geïnformeerd voelen over hun genetische vader, blijken ze meer geneigd tot rigiditeit, hetgeen volgens Luteijn c.s. (1975) betekent dat er een voorkeur bestaat voor een strak planmatige aanpak en een aversie tegen het doorbreken van gewoonten en oplossingsmethoden. Vervolgens voerden we zeven stapsgewijze multiple regressieanalyses uit, voor iedere schaal één, waarbij naast hoeveelheid en kleur van de informatie ook de in de vorige paragraaf besproken variabelen als potentiële predictoren werden betrokken. De analyses met de *inadequatie-* en de *zelfgenoegzaamheid-schaal* resulteerden in één en dezelfde significante predictor, namelijk geslacht. Op de inadequatie-schaal, waarop gespannenheid en depressiviteit naar voren komt (Luteijn et al., 1975), blijken vrouwen ietwat hoger d.w.z. inadequater te scoren dan mannen (b=0.35, $s.e.(b)$=0.17, (β=0.21), p<0.05). Op de zelfgenoegzaamheid-schaal, waarin het gaat om tevredenheid over zichzelf en een gebrek aan interesse voor anderen en hun problemen (Luteijn et al., 1975), is het effect net andersom: mannen blijken zelfgenoegzamer dan vrouwen (b=-0.66, $s.e.(b)$=0.16, (β=-0.39), p<0.001). Voor de *sociale inadequatie-schaal*, waarbij een hoge score betekent dat de persoon zich op het moment als incompetent (geremd en verlegen) beschouwt in contacten met anderen en een lage score dat de persoon vindt dat hij goed overweg kan met anderen, resulteerde de analyse in drie significante predictoren met een multiple correlatie van r=0.37. Predictoren zijn: de variabele die aangeeft of er een feitelijke ontmoeting met de genetische vader heeft plaatsgevonden (b=0.55, $s.e.(b)$=0.20, (β=0.28), p<0.01), leeftijd (b=0.03, $s.e.(b)$=0.01, (β=0.26), p<0.05) en geslacht (b=0.44, $s.e.(b)$=0.21, (β=0.21), p<0.05). Het blijkt dat geïnterviewden die hun genetische vader hebben ontmoet relatief laag scoren op sociale inadequatie, dat de scores op deze schaal toenemen met de leeftijd en dat vrouwen hoger, dat wil zeggen opnieuw inadequater, scoren dan mannen. Bij het voorspellen van *rigiditeit* werd naast de kleur van de informatie (b=-0.17, $s.e.(b)$=0.09, (β=-0.20), p<0.05), geen andere predictor gevonden. Het betreft hier dezelfde relatie als die we eerder vonden bij het berekenen van de correlaties tussen de NPV-schalen en beide informatievariabelen. Ook de analyse met de *verongelijktheid-schaal*, waarmee kritiek en wantrouwen ten aanzien van andere mensen wordt gemeten, resulteerde in één significante predictor, namelijk de variabele die aangeeft dat de vader is verdwenen na een kortstondige voor- c.q. buitenechtelijke relatie. Personen van wie de vader om deze reden ontbreekt blijken relatief hoog te scoren op verongelijktheid

(b=0.44, $s.e.(b)$=0.19, (β=0.24), p<0.05). De analyses van de schalen *Dominantie* en *Zelfwaardering* ten slotte, leverden geen significante predictoren op.

In het voorgaande werd voor drie NPV-schalen een effect van geslacht gerapporteerd, namelijk voor inadequatie, sociale inadequatie en zelfgenoegzaamheid: vrouwen maakten de indruk inadequater en minder zelfgenoegzaam te zijn dan mannen. Door Luteijn c.s. (1975) werden deze sekseverschillen niet vastgesteld. In onderzoekingen bij soortgelijke proefgroepen als de onze kwamen echter eerder ongunstige resultaten voor vrouwen naar voren. Amato en Keith (1991) bespreken een vijftal studies, waarin kinderen met een stiefvader worden vergeleken met kinderen uit één-ouder (moeder) gezinnen. Uit deze onderzoeken komt naar voren dat, waar jongens baat blijken te hebben bij een vervangende vader, bij meisjes eerder van een negatieve invloed moet worden gesproken.

Tevredenheid

Aan de geïnterviewden werd gevraagd of zij tevreden waren over hun huidige leven, werk, relaties met de andere sekse en sociale contacten. De antwoorden op deze vragen werden dichotoom gescoord: ontevreden versus tevreden. De percentages in de antwoordcategorie tevreden bedroegen respectievelijk, huidige leven: 77%, werk: 63%, relaties met de andere sekse: 56% en sociale contacten: 71%. Met behulp van logistische regressie-analyse werd nagegaan in hoeverre er relaties bestaan tussen antwoorden op de gestelde vragen enerzijds en de hoeveelheid en kleur van de informatie anderzijds. Ook hier werden de eerder besproken achtergrondvariabelen vanwege hun mogelijke invloed in de analyses betrokken. Voor de antwoorden op de vragen naar tevredenheid over het huidige leven en tevredenheid op het gebied van werk werden geen significante predictoren gevonden. De analyse met tevredenheid over *relaties met de andere sekse* resulteerde in één significante predictor, namelijk de kleur van de informatie (b=-0,43, $s.e.(b)$=0.20, p<0.05). Degenen die zich relatief positief geïnformeerd voelen, geven grotere tevredenheid aan over hun relatie *met de andere sekse*. De analyse met tevredenheid over *sociale contacten* als responsvariabele resulteerde eveneens in één significante predictor. Degenen die hun genetische vader daadwerkelijk hebben ontmoet uiten zich tevredener over hun sociale contacten dan degenen die niet over deze ervaring beschikken (b=1.17, $s.e.(b)$=0.52, p<0.05).

Hulp gezocht voor psychische problemen

De geïnterviewden kregen de vraag voorgelegd of zij ooit hulp hadden gezocht voor psychische problemen en zo ja, welke klachten zij destijds hadden. De antwoorden op de eerste vraag werden gescoord in twee categorieën: wel hulp gezocht versus geen hulp gezocht. Zevenenvijftig procent van de ondervraagden gaf aan ooit hulp gezocht te hebben voor psychische problemen, hetgeen beschouwd kan worden als een hoog percentage. Van deze groep werd door 22% relatieproblemen, 12% psychosomatische klachten, 24% depressiviteit, 21% emotionele crises en 21% andere problemen genoemd als reden voor het zoeken van hulp. Relaties tussen hulp gezocht voor psychische problemen enerzijds en hoeveelheid en kleur van de informatie en andere achtergrondvariabelen anderzijds werden onderzocht met behulp van logistische regressie-analyse. De analyses leverden geen significante predictoren op.

Gedachten over de genetische vader in het huidige leven

Voor het merendeel van de ondervraagden is het al geruime tijd geleden dat zij werden geïnformeerd over het bestaan van hun genetische vader. Met een drietal vragen werd onderzocht hoe de genetische vader de ondervraagden ten tijde van het interview nog bezighoudt. Er werd gevraagd of ze vandaag de dag nog fantasieën hebben over de genetische vader (bijv. over uiterlijk, karakter, hoe zijn leven eruit ziet), of ze denken dat ze uiterlijk op hun genetische vader lijken en of ze vinden dat ze qua karakter op hun genetische vader lijken. De antwoorden op de vragen werden dichotoom gescoord: wel fantasieën versus geen fantasieën; veel op hem lijken (uiterlijk, qua karakter) versus weinig op hem lijken. De verdeling van de antwoorden over de genoemde categorieën laat zien dat 58% van de ondervraagden aangeeft nog wel eens over de genetische vader te fantaseren, tegenover 42% niet; dat 40% het idee heeft uiterlijk sterk op de genetische vader te lijken, te-

genover 60% weinig; en dat 39% het idee heeft qua karakter sterk op de genetische vader te lijken tegenover 61% weinig. De percentages uit de 'gelijkenis qua karakter'-vraag worden interessant als we ze zetten naast vergelijkbare gegevens over de moeder en stiefvader. De antwoordpercentages op de vraag over de moeder komen dicht in de buurt van die over de genetische vader: 34% van de ondervraagden vindt dat hij of zij qua karakter sterk op de *moeder* lijkt, tegenover 66% weinig. Ten aanzien van de *stiefvader* echter, is slechts 4% van mening dat er sprake is van een sterke gelijkenis qua karakter, tegenover 96% weinig. Een vergelijking van de resultaten op de 'gelijkenis qua karakter'-vragen wekt de indruk dat een deel van de groep ondervraagden sterke erfelijkheidsopvattingen huldigt. Vervolgens voerden we per vraag een logistische regressie-analyse uit, waarbij zowel hoeveelheid en kleur van de informatie als de andere achtergrondvariabelen als potentiële predictoren werden betrokken. De analyse met fantaseren over de genetische vader leverde één significante predictor op, namelijk de variabele die aangeeft of er een ontmoeting met de vader heeft plaatsgevonden (b=-1.03, $s.e.(b)$=0.44, p<0.05). Degenen die hun vader niet in levende lijve hebben ontmoet zijn geneigd meer over hem te fantaseren. Uit de analyse met de vraag over gelijkenis qua karakter kwam dezelfde predictor naar voren: degenen die hun genetische vader hebben ontmoet blijken vaker het idee te hebben qua karakter sterk op hun genetische vader te lijken (b=1.31, $s.e.(b)$=0.46, p<0.01). De analyse met de vraag over uiterlijke gelijkenis leverde geen significante predictoren op.

Erfelijkheidsopvattingen

In hoeverre worden erfelijkheidsopvattingen door de hoeveelheid en kleur van informatie en andere variabelen bepaald? De correlaties tussen de erfelijkheidsschaal en de hoeveelheid en kleur van de informatie bedragen respectievelijk 0.04 en 0.15, beide niet significant. De daarna uitgevoerde regressie-analyse met de erfelijkheidsschaal als responsvariabele resulteerde in één significante predictor. Personen van wie de vader verdwenen is na een kortstondige relatie blijken meer belang te hechten aan erfelijke factoren dan personen bij wie de genetische vader ontbreekt door overlijden of echtscheiding (b=5.40, $s.e.(b)$=1.66, (β=0.32), p<0.005).

4 Discussie

De resultaten van het onderzoek naar de rol die informatie over de afwezige genetische vader speelt voor personen die opgroeiden met een vervangende vader kunnen als volgt worden samengevat. Met betrekking tot de *achtergronden* van informatie of kennis over de genetische vader kwam naar voren dat het gevoel goed geïnformeerd te zijn samengaat met (1) het hebben ontmoet van de genetische vader, (2) overlijden als reden van afwezigheid van de vader, en (3) op jonge leeftijd zijn ingelicht over het bestaan van de genetische vader. Voorts bleek dat de kleur van de informatie naar verhouding negatief was in het geval de afwezigheid van de vader het gevolg was van echtscheiding. Met betrekking tot de *implicaties* van geïnformeerdheid werd alleen een effect van negatief gekleurde informatie op rigiditeit en een effect van positief gekleurde informatie op tevredenheid over de relatie met de andere sekse geconstateerd. Voor het overige werden er geen effecten van de hoeveelheid en kleur van de informatie over de genetische vader gevonden, noch op persoonlijk functioneren, noch op gedachten over de genetische vader in het huidige leven, noch op erfelijkheidsopvattingen. Wèl werden invloeden vastgesteld van de reden van afwezigheid van de genetische vader en van het al dan niet hebben ontmoet van de genetische vader. In het geval de vader was verdwenen na een kortstondige relatie werd er hoger gescoord op wantrouwendheid/intolerantie (NPV) ten opzichte van anderen en was er sprake van relatief sterke erfelijkheidsopvattingen. Indien er een ontmoeting met de genetische vader had plaatsgevonden werd er relatief laag gescoord op sociale inadequatie en was er vaker sprake van tevredenheid over sociale contacten in het huidige leven. Bovendien werd er in het laatste geval minder gefantaseerd over de genetische vader en had men het idee weinig op hem te lijken qua karakter.

Opvallend is dat de geïnformeerdheid van personen een minder cruciale rol speelt dan verwacht. Eigenlijk is er een directer effect

zichtbaar geworden van wat wij vooraf benoemden als voorlopers van informatie dan van de informatie zelf: het al dan niet hebben ontmoet van de genetische vader en de reden van afwezigheid van de genetische vader blijken betere voorspellers van de huidige gedachten, opvattingen en gevoelens van de ondervraagden, dan de variabelen die aangeven hoe men zich in de loop der tijd geïnformeerd voelt over de genetische vader. Het lijkt alsof er, behalve de informatie die men heeft, nog iets anders speelt namelijk hoe bereikbaar nieuwe, aanvullende informatie is. Of men de vader ontmoet heeft of niet bleek bij onze groep ondervraagden immers alleen relevant in het geval een ontmoeting in theorie ook tot de mogelijkheden behoorde. Men lijkt zich het meest gefrustreerd te voelen als datgene, wat in principe tot de mogelijkheden behoort, niet gerealiseerd wordt.

De variable *ontmoeting met de genetische vader* is onmiskenbaar de belangrijkste in het hier besproken onderzoek. Zowel op de NPV als op de tevredenheidsvariabelen vinden we een positieve samenhang tussen ontmoeting met de vader en sociaal functioneren. Daarnaast heeft deze gebeurtenis ook gevolgen voor de rol die de vader vervolgens in de gedachten van zijn nakomeling speelt: er is minder aanleiding tot fantaseren en meer basis voor identificeren. In verband met het belang van de variabele 'ontmoeting met de genetische vader' is het een interessante vraag wat een ontmoeting nu precies aan de kennis over de genetische vader toevoegt. Het onderzoek van Triseliotis (1973) geeft aanleiding te veronderstellen dat de betekenis van een ontmoeting meer van psychologische dan informatieve aard is. De meerderheid van zijn onderzoeksgroep gaf immers aan dat zij, hoewel ze niet veel wijzer waren geworden en de gebeurtenis hen vaak teleurstelde, toch een zekere rust hadden gevonden door de ontmoeting met hun genetische ouder(s).

De *reden van afwezigheid* van de genetische vader is een volgende belangrijke variabele. De bevindingen laten zien dat het zinloos is in zijn algemeenheid uitspraken te doen over het belang van het ontbreken van de genetische vader. Het blijkt essentieel onderscheid te maken naar de oorzaak van de afwezigheid van de vader, d.w.z. of hij ontbreekt door overlijden, echtscheiding of vanwege vertrek na een kortdurende relatie. Zoals verwacht zijn voor de kinderen van overleden vaders de bevindingen het positiefst, zowel wat betreft de geïnformeerdheid over de vader als hun persoonlijk functioneren in het latere leven. Personen geboren uit een kortstondige relatie lijken naar verhouding het slechtste af. Hun latere functioneren wordt gekenmerkt door een relatief hoge mate van verongelijktheid en wantrouwen en sterke erfelijkheidsopvattingen. Kinderen van gescheiden ouders zitten hier min of meer tussenin. Een eerste voor de hand liggende verklaring voor deze verschillen is dat kinderen waarvan de vader overleden is zich naar verhouding het minst verlaten en afgewezen zullen voelen. In geval van vertrek na echtscheiding of na een kortstondige relatie is er psychologisch beschouwd sprake van verstoting door de genetische vader. De vader heeft het kind, volgens de waarneming van de laatstgenoemde, immers willens en wetens verlaten (Slijper & Verhulst, 1991; Tusenius, 1989). Voor personen, geboren uit een kortstondige relatie, komt daar nog bij dat de verbintenis waaruit zij voortkomen en daarmee in zekere zin hun hele bestaansgrond, door de omringende samenleving wordt afgekeurd. Zowel de verstoting door de vader als de afkeuring door de samenleving zou een verklaring kunnen bieden voor de verongelijktheid die werd vastgesteld op de NPV. Een verklaring voor de gunstige positie van nakomelingen van overleden vaders zou tenslotte ook nog kunnen zijn dat alleen bij deze groep de relatie met de vader in zekere zin is afgerond. In geval van overlijden ligt de zaak duidelijk: er kunnen geen stappen ten aanzien van de vader meer worden ondernomen, hetgeen een bepaalde rust geeft. In het geval van echtscheiding of kortstondige relatie is deze rust er niet. De vader is immers nog ergens, waardoor er in potentie nog van alles met de relatie kan gebeuren.

Nakomelingen van *gescheiden* ouders springen er in het onderzoek in één opzicht negatief uit: ze blijken vaker dan kinderen met een overleden vader of kinderen geboren uit een kortstondige relatie (via hun moeder) een negatief beeld te hebben opgebouwd van hun genetische vader. Hoewel deze bevinding strookt met resultaten van Hetherington (1972) en Hainline en Feig (1978), moet men er ons inziens voor oppassen deze bevinding te gene-

raliseren naar kinderen van gescheiden ouders in het algemeen. In een onderzoek in stiefgezinnen van Spruijt (1991) bleek dat in die gevallen waarin na een echtscheiding het contact met de vertrokken ouder verbroken was (hetgeen ook bij onze proefpersonen het geval is), de scheiding vaak zeer conflictueus was verlopen. Deze bevinding biedt niet alleen een verklaring voor de negatieve kleur van de informatie bij kinderen van gescheiden ouders, het maakt tevens duidelijk dat we met een specifieke subcategorie kinderen te maken hebben.

Een opvallende bevinding is dat personen die pas in de *adolescentie* van het bestaan van hun genetische vader op de hoogte werden gesteld, zich (ook op latere leeftijd) relatief het slechtst geïnformeerd voelden. Deze subjectieve beleving van een tekort aan informatie zou op drie manieren verklaard kunnen worden. Ten eerste zou de ervaring van een tekort samen kunnen hangen met een relatief grote *informatiebehoefte* op deze leeftijd. Volgens Triseliotis raken geadopteerden juist in deze periode zeer geïnteresseerd in hun adoptiestatus, ook als ze er al jaren van op de hoogte zijn. Personen die pas in of na de adolescentie te weten kwamen dat ze geadopteerd waren, hadden het gevoel niet langer te weten wie ze waren en wensten meer over hun genetische ouders te weten te komen om zichzelf opnieuw te kunnen definiëren (Triseliotis, 1973). Ook de personen uit ons onderzoek zouden vanuit een dergelijke toestand van identiteitsverwarring een dermate intense behoefte aan informatie kunnen hebben, dat de informatie die hun geboden wordt hier mager bij afsteekt. Behalve verwarring veroorzaakt late berichtgeving over genetische ouders ook een aanzienlijke deuk in het *vertrouwen* in de sociale ouders vanwege het achterhouden van informatie (Triseliotis, 1973), waarmee we zijn aangeland bij een tweede verklaringsmogelijkheid. Het beschaamde vertrouwen zou de relatie zodanig kunnen aantasten, dat men de informatie die door de sociale ouders verstrekt wordt niet zonder meer aanneemt. De schade die wordt aangericht in de relatie met de sociale ouders is waarschijnlijk sterker in de adolescentieperiode dan daarna. In deze fase zal men zich, vanwege de afhankelijkheidsrelatie die nog bestaat, meer bekocht voelen. Ten derde is er mogelijk ook sprake van een *feitelijk* informatietekort. De groep laatgeïnformeerden telt relatief veel personen die zijn geboren uit kortstondige buitenechtelijke c.q. voorechtelijke relaties. De genetische vaders uit deze relaties hebben maar kort een rol gespeeld, waardoor de moeders in de meeste gevallen ook maar weinig informatie hebben *kunnen* verstrekken. Overigens zou de verongelijktheid die bij personen geboren uit kortstondige relaties is vastgesteld in dit verband ook geïnterpreteerd kunnen worden als boosheid ten gevolge van het achterhouden van informatie door de opvoeders. Hoewel we voorop stellen dat eenduidige resultaten ontbreken, wekt deze combinatie van bevindingen de suggestie dat het weghouden van informatie een gevaar op kan leveren voor het latere functioneren.

In het onderzoek werd nadere bevestiging gezocht voor de gedachtengang dat tekortschietende informatie tot relatief sterke *erfelijkheidsopvattingen* zou leiden (Van Kampen et al., 1990). Ook met betrekking tot dit onderwerp zijn er moeilijk eenduidige conclusies te trekken. Enerzijds hebben we geen relatie gevonden tussen hoeveelheid en kleur van geïnformeerdheid en erfelijkheidsopvattingen. Anderzijds werden bij personen die geboren zijn uit een kortstondige relatie inderdaad sterke erfelijkheidsopvattingen gevonden. Als we deze bevinding zetten naast hetgeen hierboven ten aanzien van kortstondige relaties werd vastgesteld, constateren we dat de associatie tussen geboren uit een kortstondige relatie en een laat tijdstip van informeren een indicatie geeft in de richting van de erfelijkheidshypothese van Van Kampen c.s. Personen die pas tijdens de adolescentie op de hoogte zijn gesteld van het bestaan van hun genetische vader zouden juist vanwege de eerder genoemde boosheid op de moeder en vervangende vader veel toe kunnen schrijven aan de onbekende genetische vader. Aan de andere kant vonden we bij deze personen geen neiging tot fantaseren, noch tekenen van identificeren met de genetische vader op de gelijkenis-items, waardoor benadrukt moet worden dat de basis voor een conclusie ten voordele van de erfelijkheidshypothese wel erg mager is.

Uit dit onderzoek is niet naar voren gekomen dat ontoereikende informatie over de genetische vader grote gevolgen heeft voor het kind. De verschillen op het gebied van persoon-

lijkheidskenmerken en de mate van tevredenheid waarmee men later in het leven komt te staan, zijn nauwelijks opzienbarend te noemen. Toch moeten we voorzichtig zijn op deze plaats al te sterke conclusies te trekken. Een onderzoek als het onze heeft immers onvermijdelijk zijn beperkingen. Zo is er gebruik gemaakt van een specifieke operationalisatie van geïnformeerdheid. Er werd bijvoorbeeld niet verdisconteerd welke *feitelijke* informatie de proefpersonen ter beschikking hadden. En daarnaast zijn er ook allerlei variabelen niet in het onderzoek betrokken, die wèl een rol zouden kunnen spelen. Bij Amato en Keith (1991) vinden we in dit verband de volgende variabelen: de kwaliteit van de relatie met de vervangende vader (inclusief een eventueel verschillende invloed voor de seksen), conflicten binnen het gezin tussen moeder en vervangende vader, de levensstandaard van het gezin en ten slotte ook, voor die nakomelingen waarvan de genetische vader nog in leven is, de intensiteit en kwaliteit van de relatie met deze afwezige vader.

Ten slotte moet nog worden opgemerkt dat voor de groep als geheel het vastgestelde percentage personen dat hulp heeft gezocht voor psychische problemen erg hoog is, namelijk 57%. Nu is het enerzijds zeer wel mogelijk dat er een selectie-effect is opgetreden. Het is niet onwaarschijnlijk dat een groot aantal mensen zich vooral voor het onderzoek heeft aangemeld vanuit een behoefte om min of meer anoniem het een en ander van zich af te kunnen praten, waardoor personen met problemen mogelijk zijn oververtegenwoordigd. Aan de andere kant kan, mede met het oog op het feit dat bij alle geïnterviewden sprake was van een normale opvoedingssituatie, ook niet worden uitgesloten dat afwezigheid van de genetische vader tot meer problemen aanleiding geeft dan we met dit onderzoek hebben kunnen vaststellen.

Literatuur

Amato, P. R. (1987). Family processes in one-parent, stepparent, and intact families: The child's point of view. *Journal of Marriage and the Family, 49*, 327-337.

Amato, P. R., & Keith, B. (1991). Parental divorce and the well-being of children: a meta-analysis. *Psychological Bulletin, 110*, 26-46.

Bach, G. R. (1946). Father fantasies and father-typing in father-separated children. *Child Development, 17* (1/2), 63-80.

Bierkens, P. (1992). Spermadonorschap en geheimhouding. *De Psycholoog, 27*, 57-61.

Goebel, B., & Lott, S. L. (1986). *Adoptees' resolution of the adolescent identity crisis: Where are the taproots?* Paper presented at the meeting of the American Psychological Association, Washington D C

Hainline, L., & Feig, E. (1978). The correlates of childhood father absence in college-aged women. *Child Development, 49*, 37-42.

Hetherington, E. M. (1972). Effects of father absence on personality development in adolescent daughters. *Developmental Psychology, 7*, 313-326.

Hoksbergen, R. (1992). Afstammingsrust: een lezing gehouden voor genealogen in februari 1991. In *Jaarverslag van het Centraal Bureau voor Genealogie.*

Hoksbergen, R. A. C. (1987). Adolescent en geadopteerd of geadopteerd en adolescent? In R. A. C. Hoksbergen e.a. (red.), *Adolescenten in vele gedaanten: Bijdragen over adolescentie en adolescentenzorg*. Lisse: Swets & Zeitlinger.

Hoksbergen, R. A. C. (1990, 22 maart). Interview in *De Volkskrant*.

Hoksbergen, R. A. C., & Walenkamp, A. (1983), *Adoptie uit de kinderschoenen*. Meppel: Boom.

Kampen, L. van, Koops, W., Meerum Terwogt, M., & Reijnders, C. (1990). Onbekendheid met de biologische vader als een belemmerende factor in de identiteitsontwikkeling: een empirische exploratie. *Nederlands Tijdschrift voor de Psychologie, 45*, 283-288.

Luteijn, F., Starren, J., & Dijk, H. van (1975). *Handleiding bij de NPV*. Lisse: Swets & Zeitlinger.

Meerum Terwogt, M. (1991). De proefballon van Hirsch Ballin: een pleidooi voor het opheffen van de donoranonimiteit. *De Psycholoog, 26*, 396-397.

Meerum Terwogt, M. (1992). Geheimhouding? Waarom? *De Psycholoog, 27*, 62-63.

Meerum Terwogt, M., Hoeksma, J. B., & Koops, W. (1991). Opvattingen over het erfelijk karakter van persoonlijkheidskenmerken. *Nederlands Tijdschrift voor de Psychologie, 46*, 344-346.

Oshman, H. P. (1975). Some effects of father-absence upon the psychosocial development of male and female late-adolescents: Theoretical and empirical considerations. *Dissertation Abstracts International, 36*, 919B-920B.

Raad voor het Jeugdbeleid (1988). *Ouderschap zonder onderscheid: een beleidsadvies over de belangen van kinderen bij verantwoord ouderschap en kunstmatige voortplanting*. Rijswijk: Ministerie van Welzijn, Volksgezondheid en Cultuur.

Sibbing-Willems, L., & Temmink-Spits, M. (1991). Hulp aan geadopteerden bij het zoeken naar hun oorsprong. In R. Hoksbergen & H. Walenkamp (red.), *Kind van andere ouders: Theorie en praktijk van adoptie*. Houten: Bohn Stafleu Van Loghum.

Slijper, F. M. E., & Verhulst, F. C. (1991). Rapport betreffende de onderzoekbaarheid van psychische gevolgen voor KID-kinderen om geen toegang tot hun afstammingsgegevens te hebben. Rotterdam: Erasmus Universiteit.

Spruijt, E. (1991). Stiefgezinnen en stiefouderadoptie. In R. Hoksbergen & H. Walenkamp (red.), *Kind van andere ouders: Theorie en praktijk van adoptie*. Houten: Bohn Stafleu Van Loghum.

Stein, L. A., & Hoopes, J. L. (1985). *Identity formation in the adopted adolescent*. New York: Child Welfare League of America.

Stevenson, M. R., & Black, K. N. (1988). Paternal absence and sex-role development: A meta-analysis. *Child Development, 59*, 793-814.

Triseliotis, J. (1973). *In search of origins: the experiences of adopted people*. London: Routledge & Kegan Paul.

Tusenius, A. A. (1989). De gevolgen van echtscheiding voor kinderen. *Tijdschrift voor Psychiatrie, 31*, 15-28.

Verhulst, F. C., & Versluis-den Bieman, H. J. M. (1989). *Buitenlandse adoptiekinderen: Vaardigheden en probleemgedrag*. Assen/Maastricht: Van Gorcum.

Manuscript aanvaard 29-7-1993

Auteurs

H. M. Y. Koomen is universitair docent aan de sectie Orthopedagogiek van de Vrije Universiteit.

Adres: Sectie Orthopedagogiek, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT Amsterdam

J. B. Hoeksma is universitair docent aan de vakgroep Kinder- en Jeugdpsychologie van de Vrije Universiteit.

M. Meerum Terwogt is universitair hoofddocent aan de vakgroep Kinder- en Jeugdpsychologie van de Vrije Universiteit.

Adres: Vakgroep Kinder- en Jeugdpsychologie, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT Amsterdam

Abstract

The importance of knowledge about an absent genetic father

H. M. Y. Koomen, J. B. Hoeksma & M. Meerum Terwogt. Pedagogische Studiën, 1993, 70, 368-381.

This study is based on structured interviews among 102 adult subjects raised by their genetic mother and a substitute father. The questions addressed the subject's knowledge about their absent genetic fathers and their personal well-being and social adjustment. In addition, two standardized questionnaires (a personality test and a test of people's beliefs in heridity) were administered. The main results were: (1) the amount and valency of knowledge individuals claim to have plays a less prominent role in their psychosocial adjustment than was expected; (2) having met the genetic father appears to have a favourable effect on knowledge about the father as well as later social adjustment; (3) the reason of absence of the genetic father appears to be of vital importance: adult children who were born in a short-lived relationship are worse off than adult children who lost their fathers through death or divorce, both with regard to knowledge and psychosocial adjustment. Children of divorced parents have a position somewhere in the middle.

De autonomievergroting van scholen. Zin en onzin

A. M. P. Knoers

1 Omschrijving van het begrip autonomie en autonome school

Vooral nu er een akkoord schijnt te zijn gesloten tussen de minister en de onderwijskoepels over de bekostiging van autonome basisscholen, lijkt het een goed moment om over zin en onzin van autonome scholen te discussiëren. Maar dan moet wel duidelijk zijn wat onder de term wordt verstaan. Van oorsprong is autonomie een juridisch begrip uit het volkerenrecht en uit het staatsrecht. In het laatstgenoemde recht is het volgens Van Dale (11e uitg. 1984) "de bevoegdheid van rechtsgemeenschappen van lagere orde dan de staat algemeen bindende voorschriften te geven in eigen aangelegenheden". Aanvankelijk had het betrekking op provincies en gemeenten, maar later werd het ook gebruikt voor andere publiekrechtelijke rechtspersonen, zoals openbare scholen, en eveneens voor privaatrechtelijke rechtspersonen zoals bijzondere scholen.

De term 'autonome school' voor zulk een rechtspersoon 'van lagere orde' is vrij recent. De oudst mij bekende vermelding vind ik bij Langeveld (1967), waar de term een voor ons ongebruikelijke inhoud heeft, namelijk van een school wier taak "niet door de ouders als zodanig bepaald wordt of kàn worden" (l.c. p. 13). Zij heeft daar dus een omgekeerde betekenis als wij eraan toekennen en die ik voor het eerst geformuleerd vond in een artikel van Leune (1970) in dit tijdschrift, waar hij het aldus formuleert: "Autonomie heeft betrekking op de mate waarin de samenleving (hogere of lagere overheden, ouders of andere belangengroepen) zeggenschap heeft in het interne schoolleven" (l.c. p. 308). Leune wijst er daar ook op dat, hoewel voorstanders van de autonome school pleiten voor een grotere vrijheid, totale autonomie van de school geen realistische verwachting is.

Hier wordt de nieuwe term terecht verbonden met een zaak die zelf van minder recente datum is. Autonomie van scholen is in wezen een uitvloeisel van de grondwettelijke vrijheid van onderwijs, gewaarborgd door de Grondwet van 1917 (gewijzigd in 1983 en 1987), thans art. 23 en met name lid 2 daarvan, en voor wat de verbinding met de eisen van deugdelijkheid betreft ook lid 5. Maar Leune maakt terecht ook een toespitsing: de autonomie is niet totaal, zij is zoals de N.K.S.R.(1979) het formuleert in de BARAS-nota, die de autonomie eveneens verbindt met de grondwettelijke vrijheid van richting, stichting en inrichting van scholen, een *relatieve* autonomie.

De N.K.S.R. kiest in deze nota voor één van drie autonomiemodellen in de dissertatie van Marx (1975) en wel voor het model dat in een latere publikatie van Marx (1990) het *operationeel autonome* model wordt genoemd, dat hij daar onderscheidt van het beheersmatig autonome model en het strategisch autonomiemodel.

Het is dat operationele model dat het dichtst staat bij de omschrijving in "De school op weg naar 2000" (Min. v. O&W, 1988), welke ministeriële nota, evenals de "HOAK"-nota (Min.v.O&W, 1985) dit deed voor instellingen van hoger onderwijs, de gedachten in het onderwijsbeleid en in het onderwijsoverleg heeft bepaald.

Die nota schetst een "ideaalbeeld" van de autonome school waaruit het beeld van de "gewenste sturing van het onderwijssysteem" zou voortvloeien (l.c., p. 3).

De autonome school wordt in die nota niet gedefinieerd, maar in haar kenmerken of in haar functioneren omschreven.

Daarop wordt in de volgende paragraaf nader ingegaan. Terloops zij vermeld dat dit type school ook model staat voor de gedachten over de ideale school in het rapport van de Commissie Toekomst Leraarschap (1993) - naar de voorzitter mw. A. van Es ook wel de Commissie-Van Es genoemd -, waar die autonome school wordt omschreven als een moderne, flexibele professionele arbeidsorganisatie, in welk laatste begrip de commissie een toespit-

sing geeft die meer arbeidsorganisatorisch dan onderwijskundig is.

2 Kenmerken van de autonome school

Omdat daarin consequent de organisatiecomponent (a), de schoolcultuur (b), het beleidsvoerend vermogen (c) en de belangrijkste managementcriteria (d) worden verbonden, die samen aangeven hoe van de grotere beleidsruimte van de autonome school gebruik wordt gemaakt, wordt hier uitgegaan van de benadering van Marx (1990). Daarna zal worden bezien of en zo ja in hoeverre deze verschilt van het ministeriële model.

In dit model vindt een combinatie plaats van beheersmatige deregulering (lump sum) en globalisering van de onderwijskundig-organisatorische regelgeving, die de grotere beleidsruimte mogelijk maken. Marx noemt als kenmerken:

- "coöperatief gereguleerde organisatie (a), professioneel en tevens collegiaal gekleurde cultuur (b),
  integrerend beleidsvoerend vermogen, oriëntatie op professionele markt (c),
  doeltreffend en responsief in vakprofessionele en onderwijskundige zin" (d) (l.c., p. 18)

Het model van de ministeriële nota heeft veel van de genoemde kenmerken, maar verschilt daarvan toch in die zin dat het een aangescherpt operationeel model genoemd zou kunnen worden. Er ligt meer nadruk op de beheersmatige en organisatorische componenten dan op de onderwijskundige component, die door Marx veel meer wordt geaccentueerd.

Het coöperatieve in de organisatie en het collegiale in de professionele cultuur wordt in de ministeriële nota ondergewaardeerd ten koste van "een professionele leiding, goed in managementtechnieken, onderwijskundige ontwikkelingen, personeelsbeleid en financiëel en materiëel beheer" (l.c., p. 3).

Dit beleid wordt door de schoolleiding ontwikkeld, zij het samen met "verschillende functionarissen uit het onderwijzend en onderwijsondersteunend personeel". Niet duidelijk wordt of hier slechts op het middenmanagement wordt gedoeld, maar dat aan een collegiale cultuur weinig wordt gedacht blijkt uit de nadruk die de schoolleiding krijgt toegewezen in het personeelsbeleid, dus in het benoemings- en afvloeiïngsbeleid en in het promotiebeleid, maar ook in het na- en bijscholingsbeleid, hoewel de nota daarin vager blijft dan de commissie-Van Es, die vindt dat dit alles behoort tot de "professionele taak van de directie" (1993,p. 78 e.v.).

Dat de beheersmatige component in de wat vage constructie van de ministeriële nota toch is aangescherpt blijkt daaruit dat een "duidelijke lijn" wordt gezien van vergrote autonomie en financiële verantwoordelijkheid "naar de kwaliteit van het door de scholen aangeboden onderwijs" (l.c. p. 4). Dat die lijn er is wordt nergens door onderzoek aangetoond. Er worden verwachtingen geuit: "ze kunnen beter op kwaliteit sturen", en machtswoorden gesproken: "zij moeten dat ook doen" (l.c. p. 5). Dat ziet er allemaal niet zo stevig uit. Om beter zicht te krijgen op zin en onzin van de autonome school doen wij er dus beter aan niet van dit 'ideaalbeeld' uit te gaan, maar van het beter doordachte operationele model van Marx (1990).

3 Mogelijk positieve kanten van de autonomie (de zin)

De term 'mogelijk' in de titel van deze paragraaf duidt erop dat het ook hier om veronderstellingen gaat, omdat empirische gegevens niet voor handen zijn.

Het boven beschreven model van Marx (1990) kan operationeel worden genoemd, als het geschikt is om er op zodanige wijze mee te werken dat aannemelijk kan worden gemaakt dat zulke autonome scholen in belangrijke opzichten, in het bijzonder in onderwijskundige zin, beter functioneren dan de huidige, niet autonome scholen. Leune (1970) spreekt hier van 'funkties'. 'Beter functioneren' wordt hier dan verstaan in de zin van *effectiever* zijn. De school zou dan een effectieve school moeten zijn, waarvan de empirie in elk geval heeft opgeleverd dat 'vermoedelijke determinanten' van effectiviteit zijn: sterk onderwijskundig leiderschap, een veilig en ordelijk klimaat op school, nadruk op leren van basisvaardigheden, hoge verwachtingen van leerlingprestaties en

vorderingenevaluatie (Scheerens, 1989, p. 134).

Als die kenmerken in autonome scholen worden gevonden of aannemelijk gemaakt is dat mooi. Maar dat betekent nog geenszins dat scholen ten einde effectief te zijn, ook autonome scholen moeten zijn.

Veel modellen van autonome scholen benadrukken het belang van een sterk onderwijskundig leiderschap. Marx (1990) verstaat het sterke beleidsvoerend vermogen als een 'integrerend' vermogen, d.w.z. dat het management een beleid voert waarin de samenhang van het onderwijskundige, het beheersmatige en het personeelsbeleid in het oog wordt gehouden. Het lijkt aannemelijk dat zulk een beleid een belangrijke determinant van effectiviteit mag worden genoemd. Ook zullen de coöperatieve organisatie en collegiale cultuur, indien ingekaderd in het zojuist vermelde schoolleidingsbeleid, het klimaat van de school veilig en ordelijk doen zijn. Dat coöperatief gereguleerde management stelt de school ook "in staat tot marktgerichte beleidsvoering met name waar doelgroepen ook zelf een professionele cultuur hebben, zoals in het vervolgonderwijs, beroepsverenigingen en sommige sectoren van bedrijven en instellingen" (Marx, 1990, p. 19). De gerichtheid op hoge leerlingprestaties is daarin derhalve geïmpliceerd. De 'systematische intercollegiale samenwerking en kwaliteitsbewaking' die Marx (1990) aan dit model toekent zegt weliswaar meer over het onderwijs dan over de leerlingenvorderingen. Maar het ligt voor de hand of het zou althans zo moeten zijn dat goede leerlingvorderingen, periodiek gepeild, een uiterst belangrijke indicatie van kwalitatief goed onderwijs zijn.

Maar hier raken wij tevens aan de grenzen van de mogelijkheden van de autonomie. De garantie van een adequaat beleidsvoerend vermogen ontstaat niet door de grotere beleidsruimte en de lump sum-bekostiging alleen. De geboden mogelijkheden moeten ook goed gebruikt worden en dus beantwoorden aan een aantal belangrijke voorwaarden: na- en bijscholing van schoolleiders, vakprofessionele en onderwijskundige na- en bijscholing van docenten, aanwezigheid van goede leer- en toetsmiddelen, professionele interne begeleiding. Dat kan eventueel gerealiseerd worden, maar alleen indien men schaalvergroting van scholen accepteert, en dan blijft het nog de vraag of de kosten niet hoger zijn dan de baten en of de kosten adequaat worden aangewend, niet alleen in het beperkte belang van de school, maar vooral in dat van de leerlingen, om wie het per slot van rekening toch dient te gaan.

Daarmee komen wij dan tot de bezwaren tegen de autonome school die in de volgende paragraaf worden bediscussieerd.

Op één positief neveneffect van de lump sum-bekostiging van scholen, één der pijlers van de autonomie, mogelijk geworden ten gevolge van het in de aanhef van deze discussiebijdrage genoemde akkoord tussen minister en onderwijskoepels, dient nog gewezen te worden. Als de bepleite nieuwe juridische mogelijkheid om openbare scholen een apart bestuur te geven, los van de gemeente, ook politiek haalbaar blijkt, vervalt daarmee de 'overschrijdingsregeling' ten gevolge waarvan de gemeente alle bedragen besteed aan nodige verbetering van openbare scholen ook aan de bijzondere scholen in de gemeente moet uitbetalen. Dan zal slechts het principe gelden van gelijke behandeling in gelijke omstandigheden, hetgeen rechtvaardiger èn goedkoper is.

4 Mogelijk negatieve kanten van de autonomie (de onzin)

Leune (1970) spreekt hier van 'dysfunktie', hetgeen hij evenmin als de term 'funktie' een normatieve lading geeft.

In het voorafgaande mag zijn gebleken dat de positieve kanten van de autonome school eventueel slechts oplichten, als de school erin slaagt alle mooie voornemens in reëel schoolbeleid om te zetten. Maar zolang het tegendeel niet is gebleken kan op grond van de samenhang in een operationeel autonoom beleid, zoals hierboven is geschetst, aan de autonome school de 'benefit of the doubt' worden gegeven.

Maar de twijfel is niet weg. In een aantal opzichten blijft die zelfs bovendrijven.

Zo ziet Leune (1970) in de autonome school wel de interne democratisering bevorderd worden. Bij coöperatieve organisatie en collegiaal gekleurde schoolcultuur kan men zich daarbij iets voorstellen. Hoewel de "gezonde rivaliteit met andere scholen" van de ministeriële nota

(l.c. p. 5) er ook toe kan leiden dat de cultuur onder leraren heel professioneel en collegiaal is zonder dat de leerlingen echt bij de beleidsvoering betrokken worden. En uit onderzoek blijkt dat het participeren en functioneren van leerlingen in medezeggenschapsraden, als zij daar al worden toegelaten, niet optimaal is (Smit, Reymer, Vrieze, Ruijs, 1987).

Tegenover de mogelijke positieve effecten van de autonomie op de interne democratisering staat dat de externe democratisering daardoor volgens Leune gevaar kan lopen. Bekend is de weinig positieve verdeling van 1.90 leerlingen in het OVB over de scholen van verschillende grondslag, de meeste van deze leerlingen komen voor in openbare scholen (Jungbluth, Verhaak, Vrieze, 1989).

Volgens deze onderzoekers vraagt het vraagstuk van onderwijs-ongelijkheid "om landelijk beleid. Decentralisatie is slechts verantwoord in zoverre de huidige, relatief hoge prioritering van dit vraagstuk niet kan worden doorkruist door andere prioriteringen door lagere overheden" (l.c., p. 9). En dat blijkt bij het vrijlaten van de invulling van het beleid, hetgeen door de vergroting van de autonomie van scholen zeker zal toenemen, nu al het geval te zijn waar de invulling voor wat het OVB betreft gewoonlijk neerkomt op verkleining van de groepen (Jungbluth, e.a., 1989, p. 65).

Wat de autonome gebruikmaking van bijen nascholingsgelden betreft heb ik eerder de vrees geuit," dat de kortzichtige eigen belangen van scholen de professionalisering van leraren en de kwaliteit van onderwijs niet werkelijk zullen bevorderen" (Knoers, 1992, p. 43). In antwoord op gestelde kamervragen over die uitspraak heeft de toenmalige staatssecretaris van O. en W., Wallage, gesuggereerd dat ik over autonome scholen bevoogdend oordeelde en de Kamer heeft dat geslikt. Maar hij vergat erbij te melden dat mijn vrees, althans voor wat het basisonderwijs betreft, door onderzoek wordt gesteund. Van Tulder (1992) constateert namelijk in haar onderzoek dat "scholen, nascholing eerder op een ad hoc basis bezien dan in een beleidskader-op-langere-termijn plaatsen" (l.c., p. 151).

En mijn vrees wordt bepaald niet verkleind door de WRR-voorstellen (1991) die ook de middelen van schoolgerichte onderwijsverzorging alsmede de vernieuwingsgelden aan de autonome scholen toekennen. Na de vernietigende kritiek van Slavenburg (1992) op genoemd WRR-advies in dit tijdschrift acht ik mij van een nadere discussie daarover ontslagen. Wat in jaren aan landelijk innovatiebeleid is opgebouwd dreigt door de vraagfinanciering van nascholing, begeleiding en innovatie in belangrijke mate verloren te gaan. En dat zou echt onzinnige autonomievergroting zijn, te meer waar af te wijzen commercialisering (Janssen, 1991) dreigt, wanneer niet degenen die vanouds de expertise hebben verworven, maar ook handige 'ondernemende' figuren zich op de nascholings-, begeleidings- en vernieuwingsmarkt gaan begeven. En gezien de explosieve groei van de zelfstandige adviseursmarkt is de vrees daarvoor zeer reëel te noemen. Dat de aanbodfinanciering beëindigd gaat worden is na de brief van 26.03.1993 van de minister-president[1] wel bijna zeker.

Naarmate scholen groter zijn, zullen zij meer interne expertise kunnen opbouwen, maar die toch alleen in stand kunnen houden door voortdurend contact te houden met het wetenschappelijk onderzoek dat in universiteiten, hogescholen en para-universitaire onderzoeksinstituten wordt verricht en door ondersteuningsinstellingen wordt geïmplementeerd. Scholen zijn instellingen waar onderwijs wordt gegeven; onderzoek en ontwikkelingswerk alsmede de begeleiding van moeilijke problemen moeten zij aan andere instellingen overlaten en ook alleen in goede afspraken met die instellingen komen tot de bepaling wat voor de eigen school zinnig is en wat de vernieuwing niet belemmert.

Kleine scholen zullen in een veel grotere afhankelijkheidsrelatie blijven en ook hun autonome krachten vooral moeten besteden aan goed overleg over de door hen te besteden middelen.

5 Kort besluit

Wat in de vorige paragraaf vooral is beklemtoond is dat de autonomie (de toegekende vrijheid) van scholen de deugdelijkheid van onderwijs niet in gevaar mag brengen.

De eisen van deugdelijkheid worden volgens art. 23, lid 5 van de Grondwet bij de wet geregeld. Dat kan uiteraard een materiële wet

zijn als een A.M.v.B., maar het kan niet aan het vrije spel van krachten, autonome scholen en vrije onderwijsondernemers worden overgelaten. Beroepsvereisten van leraren - nu ook al weer ter discussie in de commissie-Van Es - en kwaliteitsbewaking achteraf zijn onvoldoende, omdat onderweg, in het onderwijsproces, van alles mis kan gaan, terwijl professionele inzet van begeleiders en ontwikkelaars kan voorkomen dat achteraf van alles moet worden gerepareerd. Dat betekent niet dat aan scholen niet van alles te verbeteren valt, o.a. door de maatregelen die een vergroting van de autonomie vereisen, en dat van onderzoekers, ontwikkelaars en begeleiders niet gevraagd mag worden zich nog eens grondig te bezinnen op hun interactie met scholen die zeker nog niet optimaal is.

Een nauwkeurige afweging van zin en onzin van autonomie leidt m.i. echter voorlopig tot de conclusie dat aan het adjectief 'relatief' zeker niet minder aandacht moet worden geschonken dan aan het substantief 'autonomie'.

Noot

1 *Onderwijsverzorging in de toekomst.* Brief van de Minister-President, Minister van Algemene Zaken aan de voorzitter van de Tweede Kamer. 's-Gravenhage: Sdu.

Literatuur

Commissie Toekomst Leraarschap (1993). *Een beroep met perspectief. De toekomst van het leraarschap.* Leiden: CTL/DOP.

Janssen, G. (1991). *Onderwijsverzorging. Wie betaalt en wie bepaalt.* 's-Hertogenbosch: KPC.

Jungbluth, P., Verhaak Ch., Vrieze, G. (1989). *Beleidsadvies formatieregeling en onderwijsvoorrang.* Nijmegen: ITS.

Knoers, A. M. P. (1992). De leraar centraal. In H. J. M. Hoefnagel (red.), *Een goede raad, de leraar tussen onderwijs en arbeidsmarkt.* 's-Gravenhage: Onderwijsraad.

Langeveld, M. J. (1967). *Scholen maken mensen.* Purmerend: Muusses.

Leune, J. M. G. (1970). Funkties en dysfunkties van de autonome school. *Pedagogische Studiën, 49,* 308 - 314.

Marx, E. C. H. (1975). *De organisatie van scholengemeenschappen in onderwijskundige optiek.* Groningen: Tjeenk Willink.

Marx, E. C. H. (1990). *Schoolontwikkeling en Milieugericht Onderwijs.* Alphen a/d Rijn: Samsom / H. D. Tjeenk Willink.

Ministerie van Onderwijs en Wetenschappen (1985). *Hoger Onderwijs: autonomie en kwaliteit. Beleidsnota.* Zoetermeer: Ministerie van Onderwijs en Wetenschappen.

Ministerie van Onderwijs en Wetenschappen (1988). *De school op weg naar 2000. Een besturingsfilosofie voor de negentiger jaren.* Zoetermeer: Ministerie van Onderwijs en Wetenschappen.

N.K.S.R. (1979). *Bouwen aan de relatief autonome school. Diskussienota.* 's-Gravenhage: N.K.S.R.

Scheerens, J. (1989). *Wat maakt scholen effectief?* 's-Gravenhage: SVO.

Slavenburg, J. H. (1992). De onderwijsverzorging in de toekomst. Een samenvatting met commentaar op het advies van de WRR. *Pedagogische Studiën, 69,* 133 - 143.

Smit, F., Reymer, G., Vrieze, G., Ruijs, A. (1987). *Het functioneren van de medezeggenschap in het onderwijs.* Nijmegen: ITS.

Tulder, N. A. Th. M. van (1992). *Nascholing en onderwijsvernieuwing.* Academisch proefschrift, Katholieke Universiteit, Nijmegen.

Wetenschappelijke Raad voor het Regeringsbeleid (1991). *De onderwijsverzorging in de toekomst.* Rapporten aan de Regering nr. 40.'s-Gravenhage: SDU.

Manuscript aanvaard 30-8-1993

Auteur

A. M. P. Knoers was als hoogleraar onderwijskunde ten behoeve van de Universitaire Lerarenopleiding verbonden aan de Katholieke Universiteit te Nijmegen en was voorzitter van de Onderwijsraad.

Adres: Kameelstraat 19, 6531 RN Nijmegen.

De Onderwijsresearchdagen 1993 in Maastricht: 26-28 mei

Inleiding (P. J. J. Stijnen, Open Universiteit, Heerlen)

Van 26 tot en met 28 mei 1993 vonden de traditionele Onderwijsresearchdagen (ORD) plaats. Deze keer in Maastricht georganiseerd door de vakgroep Onderwijsontwikkeling & Onderwijsresearch van de Rijksuniversiteit Limburg.

Ook nu weer was er grote belangstelling: het aantal inschrijvingen bedroeg 475. Van de zijde van potentiële paperlezers was er een bijzonder groot aanbod aan abstract-voorstellen (n=300). Dit heeft er toe geleid dat de themacoördinatoren voor de onaangename opdracht stonden een aantal abstract-voorstellen voor papers op grond van kwantitatieve criteria te weigeren (n=100).

De presentaties tijdens de ORD gaven een grote variatie te zien: 'klassieke' voordrachten, thematische paper- en postersessies, symposia, forumdiscussies, ronde-tafel- en consultatiegesprekken en plenaire voordrachten. Voor de meeste thema's was een key-note spreker uitgenodigd. Voorafgaand aan de ORD werden de *Proceedings* verstrekt (Van Berkel, 1993).

Op woensdagavond 26 mei vond de openingszitting plaats. Achtereenvolgens werd het woord gevoerd door H. van Berkel, de voorzitter van de organisatiecommissie, de Maastrichtse rector-magnificus J. Cohen en de voorzitter van de Vereniging voor Onderwijsresearch (VOR) N. Lagerweij (over onderwijsonderzoek anno 1993, over de positie van onderwijskunde als studierichting, over de verbreding van de VOR; zie *Tijdschrift voor Onderwijsresearch*, 1993, *18*(4). Tevens werd door prof. dr. B. Creemers, voorzitter van de jury, de VOR-prijs voor de beste dissertatie uitgereikt aan J. Vermunt van de KUB. De openingslezing over de verwevenheid van onderwijs en onderzoek ('*Schering en inslag van hetzelfde weefsel*') werd gehouden door H. Crombach.

Voor het eerst waren de Onderwijsresearchdagen gestructureerd volgens de nieuwe divisiestructuur (thema's) van de VOR: Curriculum & Organisatie, Onderwijs & Samenleving, Lerarenopleiding & Leraarsgedrag, Hoger Onderwijs & Volwasseneneducatie, Leren & Instructie, Methodologie & Evaluatie en Beroeps- & Bedrijfsopleidingen. Er was ook een thema toegevoegd vanuit de plaatselijke organisatie: Probleemgestuurd Onderwijs.

In deze kroniek zijn de impressies van een aantal personen gegroepeerd rondom de genoemde acht thema's. Het gaat nadrukkelijk om impressies waarbij de selectie een persoonlijke is en de oordelen voor rekening van de auteurs komen.

Van den Akker (UT) gaat in op 'Curriculum en Organisatie'. Hij wijst er onder meer op dat dit thema in acht aspecten werd uitgesplitst en dat het een wijs besluit lijkt in de toekomst uit te gaan van twee afzonderlijke thema's, namelijk 'Curriculum' en 'Beleid en organisatie in het Onderwijs'.

Bosker (UT) stelt 'Onderwijs en Samenleving' aan de orde. Hij constateert onder meer een goede kwaliteit aan papers, zowel methodologisch als theoretisch. Verder ziet hij een verschuiving in de doelgroepen in de richting van allochtonen, waarmee ook weer een inhoudelijke verschuiving optreedt: getracht wordt vroeger dan voorheen (vanaf het vierde levensjaar) de mechanismen voor ongelijkheid op te sporen.

G. ten Dam, G. Rijlaarsdam en J. Terwel (UvA) leveren een bijdrage naar aanleiding van het thema 'Lerarenopleiding en Leraarsgedrag'. Zij signaleren onder meer als 'zorgpunt' dat er binnen 'hun' thema weinig wordt voortgebouwd op voorafgaand onderzoek. Indien dit een algemene tendens voor de onderwijsresearch zou zijn dan is dit een bijzonder ernstig te nemen vaststelling.

'Hoger Onderwijs en Volwasseneneducatie' vormt het onderwerp van bespreking door M. Weide (RUG). Zij stelt drie actuele onderwerpen aan de orde: onderwijs-op-maat, studierendement en kwaliteitszorg.

J. van der Linden (UU) geeft impressies over het thema 'Leren en Instructie'. Hij benadrukt onder meer het procesgerichte karakter van de instructie, het belang van cognitieve strategieën bij leerlingen, de betekenis van so-

ciale interactie en het feit dat leren de gehele persoon betreft en niet alleen het 'verstand'.

De tekst van P. Sanders (CITO) heeft het thema 'Methodologie en Evaluatie' tot onderwerp. Hij wijst er onder andere op dat een groot aantal papersessies betrekking had op theoretische en praktische aspecten van studietoetsen en andere meetinstrumenten.

J. van der Sanden (KUB) bespreekt het thema 'Beroeps- en Bedrijfsopleidingen'. Hij gaat in op de key-note lezing door prof. Mandl, de drie symposia, de gepresenteerde papers en de forumdiscussie. Hij rapporteert onder meer dat uit een van de gepresenteerde onderzoeken bleek dat het karakter van in samenwerking met het bedrijfsleven gelegitimeerde eindtermen ingrijpend verandert tijdens het proces van curriculumontwikkeling: het uiteindelijk ontwikkelde lesmateriaal vertoonde weinig sporen van de oorspronkelijke (wenselijk geachte) doelstellingen.

J. Willems (KUN) tot slot doet verslag van zijn ervaringen met betrekking tot het thema 'Probleemgestuurd Onderwijs'. Hij constateert - in tegenstelling tot voorafgaande jaren - een stijgend aantal voordrachten waarin nieuwe resultaten worden gepresenteerd.

De deelnemers aan de kroniek tonen zich in het algemeen tevreden over de inhoudelijke kwaliteit en de presentatievorm. Of dit ook te maken heeft met de voor het eerst gehanteerde themaindeling kan blijken in 1994 als de ORD in Utrecht zal plaatsvinden (24 tot en met 26 mei in de Jaarbeurs).

Curriculum & Organisatie (J. J. H. van den Akker, Universiteit Twente)

In het thema 'Curriculum en Organisatie' werden 21 papers gepresenteerd, die door de themacoördinatoren (waarschijnlijk met enige moeite) gegroepeerd waren in acht 'thematische papersessies' met als titels: computer en curriculum, aspecten schoolklimaat, schoolorganisatie, schoolorganisatie basisonderwijs, effecten van onderwijsbeleid, implementatie, schooleffectiviteit, curriculum en organisatie.

Opvallend in dit diffuse aanbod was het wel zeer bescheiden aantal curriculumpapers (tijdens de AERA-conferentie in april j.l. in Atlanta waren meer Nederlandse curriculumpapers te beluisteren dan op deze ORD!), alsmede het ontbreken van key-note lezingen, symposia, fora en dergelijke die in diverse andere thema's als 'aandachttrekkers' fungeerden. Inmiddels is na enige discussie binnen de VOR besloten afzonderlijke divisies te creëren voor de thema's 'Curriculum', respectievelijk 'Beleid en Organisatie in het Onderwijs'. Dat lijkt een wijs besluit dat er hopelijk toe zal bijdragen dat volgend jaar de beide divisies zich beter zullen manifesteren. Over een deel van de prestaties in het thema volgen nu enkele impressies.

Onder de noemer *'computer en curriculum'* vonden enkele afzonderlijke presentaties plaats vanuit de UT: een verkennende studie naar computerondersteunde curriculumontwikkeling; een terugblik op het 'Proefstation Oost-Nederland', toegespitst op vereisten voor de invoering van onderwijskundig computergebruik; en een onderzoek naar implementatiebevorderende kenmerken van courseware.

Onder de noemer *'implementatie'* werden een tweetal Nijmeegse papers en een Amsterdams paper gepresenteerd. Het eerste Nijmeegse paper (Van den Goor et al.) betrof een onderzoek naar de institutionalisering van een programma voor begrijpend lezen. Daarin wordt gebruik gemaakt van zgn. 'configuratievragenlijsten' voor leerkrachten om de programmarealisatie in de lespraktijk te meten. De discussie spitste zich vooral toe op de beperkingen van dergelijke zelfrapportages en hoe die eventueel te ondervangen. Uit het andere Nijmeegse paper (Raemaekers & Veenman) over de institutionalisering van teamgerichte nascholing bleek dat er sprake was van duidelijke lange termijn effecten, maar ook van een opmerkelijk gering effect van de component coaching, hetgeen aanleiding vormde tot de nodige discussie. Het Amsterdamse paper (Mellink & Nool) ging over observatiemethoden die gehanteerd werden in een vergelijkend onderzoek naar de effecten van groepswerk ten opzichte van een klassikale aanpak in het scheikunde-onderwijs. Daarbinnen werden twee observatiemethoden toegepast: een computerondersteunde aanpak om enkele algemeen-onderwijskundige lespatronen te registreren en een sterk kwalitatief geaarde werkwijze om vakdidactische aspecten te kunnen analyseren.

Binnen het subthema *'schooleffectiviteit'* presenteerde eerst Hofman (RUG), mede naar aanleiding van de reacties op haar recente proefschrift, een secundaire analyse van onderzoeksgegevens over de relaties tussen denominatie, schoolbestuur en schooleffectiviteit. Hoewel die relaties niet geheel doorzichtig zijn, lijkt de conclusie dat de invloed van denominantie op schooleffectiviteit vooral verloopt via de organisatiestructuur en -cultuur van schoolbesturen. Het functioneren van die besturen blijkt te variëren over de denominaties. Vervolgens presenteerde Knuver (UT) een onderzoek naar de relatie tussen klas- en schoolkenmerken en het affectief functioneren van leerlingen. Affectief functioneren blijkt vooral bevorderd te worden door een ordelijk en duidelijk, maar ook vriendelijk en soepel klimaat. De relatie tussen cognitief en affectief functioneren lijkt overigens zwak. Van der Werf (RION) ging tot slot in op verschillen in school- en instructiekenmerken tussen scholen die wat effectiviteit betreft stabiel hoog, gemiddeld en laag zijn. De twee hoofdconclusies waren dat (i) de stabiliteit van schooleffectiviteit gering is, en (ii) instructiekenmerken (bijvoorbeeld effectieve leertijd) er meer toe doen dan schoolkenmerken (bijvoorbeeld onderwijskundig leiderschap). Vooral die laatste conclusie bleek (opnieuw) een bron van discussie over de mogelijke verschillen tussen effectiviteitskenmerken van Nederlandse en Angelsaksische scholen.

Onder de brede noemer *'curriculum en organisatie'* vonden twee presentaties plaats. De eerste van Bos en Koelma (UT) over het gebruik en de betekenis van PPON-produkten maakte duidelijk dat de bekendheid met en zeker het benutten van PPON-gegevens nog laag is. Het tweede paper (De Groot & Witziers, UT) over marketing door scholen in het funderend onderwijs accentueerde de toegenomen strijd om een aandeel op de leerlingenmarkt en verschafte een overzicht van de middelen waarvan scholen zich daarbij (kunnen) bedienen.

Onderwijs & Samenleving (R. Bosker, Universiteit Twente)

Met 28 papers en 4 symposia vormde het thema 'Onderwijs en Samenleving' een van de qua volume grootste onderdelen van de Onderwijsresearchdagen in Maastricht. Daarmee, en vooral door de goede kwaliteit, zowel theoretisch als methodologisch, profileerde de onderwijssociologie zich prominent binnen het nationale forum van onderwijsonderzoekers. Ondoenlijk als het is in een kort bestek een inhoudelijke samenvatting te geven van de gepresenteerde lezingen, volstaan we met het constateren van enige trends.

De thema's 'ongelijkheid in het onderwijs' en 'onderwijs en arbeidsmarkt' blijven vooralsnog de centrale thema's. Binnen deze thema's treden er echter enige elkaar aanvullende veranderingen op. Sekse-ongelijkheid in het onderwijs lijkt zich meer en meer te concentreren op het einde van het voortgezet onderwijs en bij de overgang van het onderwijs naar de arbeidsmarkt. Pregnant komt uit de gepresenteerde *papers* naar voren dat pakketkeuzen in het voortgezet onderwijs het eerste struikelblok vormen voor meisjes, en dat vervolgens opleiding voor vrouwen aanzienlijk minder rendeert dan voor mannen. Aanvullende opleidingen in het tweede-kans circuit vermogen dit niet te verhelpen, zoals ook nog eens gedemonstreerd werd bij presentaties over het extern rendement van het MBO. Dat verschillen in vormingsaanbod (in casu verschillend samengestelde vakkenpakketten) hiervoor verantwoordelijk zouden kunnen zijn, lijkt een voor de hand liggende maar nog niet onderzochte hypothese.

De tweede opvallende verschuiving is die van de doelgroepen. Concentreerde het ongelijkheidsonderzoek zich oorspronkelijk op meisjes en leerlingen uit de (hand)arbeidersmilieus, langzamerhand lijkt de laatste categorie vervangen te worden door de allochtonen. Gezien het door Bakker en Cremers (CBS) gepresenteerde trendonderzoek op basis van vier cohorten leerlingen naar ongelijke kansen, is dit laatste niet zo verwonderlijk. Het lijkt alsof de ongelijkheidssituatie zich na de invoering van de mammoetwet gestabiliseerd heeft. Na 1965 verandert het effect van milieu op de keuze voor voortgezet onderwijs niet of nauwelijks meer, zij het - en dat wordt niet door Bakker en Cremers vermeld - dat het effect inmiddels pas na twee à drie jaar voortgezet onderwijs wordt vastgesteld, waar het vroeger meteen bij de entree in het voortgezet onder-

wijs werd geconstateerd. In die eerste jaren neemt de ongelijkheid dus niet verder toe. In alle sectoren en fasen van het onderwijs, zo bleek uit diverse papers, is de achterstand van allochtonen op de autochtone leerlingen pregnant.

Met de keuze van een nieuwe doelgroep, de nieuwe onderlaag in onze samenleving, verandert er echter inhoudelijk tevens iets. Zoals duidelijk naar voren komt uit de papers gepresenteerd binnen het symposium *'de leer- en opvoedingssituatie van jonge autochtone en allochtone leerlingen'* wordt getracht, vroeger dan voorheen namelijk reeds vanaf het vierde levensjaar, de mechanismen te vinden die ten grondslag liggen aan deze ongelijkheid. Daarmee komt echter direct een frustrerend fenomeen om de hoek kijken: het gezin (met name de kwaliteit en kwantiteit van de ouder-kind interacties) is doorslaggevender dan de school, maar dat gold natuurlijk al decennia lang tevens voor het probleem van de onderwijsachterstand van (hand)arbeiders leerlingen. Nieuw is wel dat met name taalinteracties nauwkeurig (in het onderzoek van Leseman c.s. (RISBO)) door observaties worden vastgesteld, en dat niet volstaan wordt met door middel van interviews verkregen informatie over opvoedingspraktijken.

Hiermee dringt zich de vraag op wat een en ander betekent voor onderzoek van ongelijkheid in het onderwijs. In de eerste plaats krijgen we hiermee nader inzicht in de grenzen van de mogelijkheden van de school het ongelijkheidsfenomeen aan te pakken (bezie in dit verband de weinig bemoedigende resultaten van de OVB-evaluatie). In de tweede plaats zal hardnekkiger dan voorheen onderzocht dienen te worden of schooleffecten (ofwel verschillen tussen scholen in prestaties van leerlingen) eigenlijk niets anders zijn dan gezinseffecten (ofwel verschillen tussen scholen zijn terug te voeren op gemiddelde verschillen in ouder-kind-interacties). Tot de verbeelding sprekend is in dit verband het empirische feit dat verschillen tussen scholen in leerlingprestaties na de zomervakantie groter zijn dan daarvoor. In de derde plaats ontstaat de behoefte de doelgroep nauwkeuriger te definiëren: is het inderdaad de nieuwe onderlaag van de samenleving of gaat het hier om een cultureel specifiek eigen groep die zich boven op haar sociale achterstand door culturele, linguïstische of andere kenmerken onderscheidt van het klassieke handarbeidersmilieu. En in hoeverre zijn deze etnische kenmerken aan verandering onderhevig als gevolg van de confrontatie met 'de' Nederlandse cultuur? Ongelijkheid in het onderwijs, hardnekkig als het fenomeen is, is voorlopig niet van de onderzoeksagenda weg te denken, evenmin als het longitudinale schoolloopbaanonderzoek dat als basis dient om het verschijnsel in kaart te brengen. Dat scholen en leerkrachten met name belangrijk zijn voor kinderen in achterstandssituaties, belangrijker dan voor de bevoorrechte leerlingen (een empirisch meer algemeen onbekend gegeven) zou bovendien aanleiding moeten zijn tot een meer integrale bestudering van het ongelijkheidsfenomeen, zoals dit nu bijvoorbeeld in het onderzoek, gepresenteerd binnen het symposium over jonge allochten, aan de orde is. In die zin zou ook een inhoudelijk andere invulling van het schooleffectiviteitsonderzoek (ooit ontstaan uit het onderzoek naar ongelijke onderwijskansen, maar zich nu richtend op de 'gemiddelde' leerling) wenselijk zijn.

Lerarenopleiding & Leraarsgedrag (G. ten Dam, G. Rijlaarsdam en J. Terwel, ILO, Universiteit van Amsterdam)

Het rapport van de commissie Van Es *'Een beroep met perspectief. De toekomst van het leraarschap'* heeft de gemoederen danig bezig gehouden. Sinds geruime tijd staat de beroepsloopbaan van docenten en de kwaliteit van de lerarenopleiding weer hoog op de politieke agenda. In een serie openbare discussiebijeenkomsten in den lande tracht de commissie Van Es te voorkomen dat haar aanbevelingen in de la verdwijnen. Eén van de discussiebijeenkomsten vond plaats tijdens de Onderwijsresearchdagen 1993, waarmee het belang werd onderstreept dat in onderzoekerskringen wordt gehecht aan de toekomst van het beroep. Niet alleen tijdens de panel-discussie werd gedebatteerd over het leraarsberoep, het werk van de commissie Van Es klonk ook door in de *key-note* lezing van Knoers (KUN).

In een doorwrocht betoog beredeneerde hij dat een grotere autonomie van scholen onder zekere voorwaarden kan bijdragen tot een grotere professionaliteit van docenten in het voort-

gezet onderwijs. In zo'n school is het mogelijk randvoorwaarden te scheppen waarin men kan groeien tot wat Knoers noemde 'extended professionalism', die onder meer gekenmerkt wordt door professionele samenwerking, belangstelling voor vakoverstijgende activiteiten, het vergelijken van de eigen praktijk met die van anderen, de instelling dat onderwijzen een rationele kunde is. Deze professionaliteit is in tegenstelling tot de zienswijze van Van Es et al. niet alleen een kwantitatieve verandering, - de docent verricht meer taken dan alleen lesgeven - maar ook een kwalitatieve verbetering: een dergelijke organisatie maakt gebruik van gesystematiseerde praktijkkennis, heeft een goed functionerend nascholingsbeleidsplan en is een leerplaats voor stages waarin docenten met een open, onderzoekende houding aspirant docenten inscholen.

Binnen het thema Lerarenopleiding en Leraarsgedrag werd voorts opvallend veel aandacht geschonken aan de *micro-situatie*: het leren van leerlingen. In een panel-discussie onder leiding van de Leidse hoogleraar Verloop debatteerden Wubbels (UU), Peters (RUG), Pieters (UT) en Simons (KUN) over de bruikbaarheid van onderwijspsychologisch onderzoek voor de lerarenopleiding. De startvraag die Verloop in het midden wierp, bleek interessante stellingen op te leveren: welke verschillen in de onderwijsleersituatie anno nu en 1970 zijn toe te schrijven aan onderwijspsychologisch onderzoek? Pieters verstoutte zich te stellen dat er van de lerarenopleiding niets meer over bleef als er geen resultaten van onderwijspsychologisch onderzoek deel uit zouden maken van de opleiding. Simons stelde dat de aandacht voor zelfstandigheidsontwikkeling, hogere cognitieve vaardigheden en transferproblemen zonder meer nieuw genoemd konden worden. Wubbels bestreed dit laconiek door de onderwerpen op te sommen waarmee hijzelf twintig jaar geleden tijdens zijn eigen lerarenopleiding werd geconfronteerd. Peters viel Wubbels bij: nog steeds zijn docenten bezig met wat goed onderwijs vraagt aan kennis, inzicht en vaardigheden. Wat de psychologie doet is er telkens andere etiketjes op te plakken. Ook de stelling van Simons dat de problemen weliswaar nog steeds die van twintig jaar geleden zijn, maar dat de psychologie tegenwoordig veel concretere antwoorden heeft, kon hen niet overtuigen. In onderwijspsychologisch onderzoek ontbreekt de vakinhoudelijke component veel te vaak en worden pedagogische waarden veronachtzaamd. De verdere discussie, waarbij ook de zaal betrokken werd, ging over de onvermijdelijke vraag wie er nu toch verantwoordelijk voor was dat er zo weinig van de onderwijspsychologie de klas in komt. Dit terwijl Nederland het meest vertakte net van didactische vertaalinstituten kent van de wereld: de pedagogische centra.

Behalve key-note lezingen en panel-discussies werd er uiteraard in het thema 'Lerarenopleiding en Leraarsgedrag' een groot aantal *onderzoeken* gepresenteerd, al dan niet geclusterd in symposia. Het merendeel van deze onderzoeken beoogt expliciet een bijdrage te leveren aan de kwaliteit van het leraarsberoep en de lerarenopleiding. We zullen de resultaten van deze onderzoeken hopelijk in de komende periode in de tijdschriften zien verschijnen.

Twee observaties naar aanleiding van de gepresenteerde onderzoeken willen we u niet onthouden.

Allereerst geven steeds meer onderzoekers dan alleen zij die zich bewegen binnen de traditie van de cultuurhistorische school, zich rekenschap van de betekenisgeving door leerlingen. Onder de populaire en voor velerlei invullingen vatbare noemer 'constructionisme', wordt enerzijds aandacht geschonken aan het proces van kennen: de zowel binnen als buiten de school opgedane ervaringen van leerlingen spelen een rol in de wijze waarop 'de werkelijkheid' door hen wordt gerepresenteerd. Anderzijds werd in verschillende onderzoekpresentaties nadrukkelijk aandacht geschonken aan de zingeving van leerlingen. Het handelen van leerkrachten moet betekenis hebben voor leerlingen wil er een leerproces (kennen, begrijpen, reflecteren) plaatsvinden. De rol van de docent verschuift daarmee van een sturende naar een meer ondersteunende.

Ten tweede bleek het onderzoek dat tijdens de ORD gepresenteerd werd doorgaans kleinschalig. Er worden relatief veel designs gehanteerd zonder controlegroepen, waardoor de getrokken conclusies discutabel zijn. Mogelijk spelen hierbij krimpende onderzoeksbudgetten een rol. Het lijkt erop dat onderzoekers die vroeger met externe budgetten langlopende en grootschalige onderzoeken konden verrichten,

nu hebben moeten uitwijken naar kleinschalig onderzoek dat door studenten en/of AIO's wordt uitgevoerd. Weinig onderzoek bouwt theoretisch en empirisch expliciet op voorafgaand onderzoek voort, al zijn er positieve uitzonderingen. Al met al vindt er echter minder kennisaccumulatie plaats en worden in het verrichte onderzoek minder nieuwe wegen verkend dan in principe mogelijk zou zijn.

Hoger Onderwijs & Volwasseneneducatie (M. G. Weide, COWOG, Rijksuniversiteit Groningen)

Het thema 'Hoger Onderwijs en de Volwasseneneducatie' omvatte zes symposia en negen losse paperpresentaties. Onderwerpen waren onder meer: interactieve leeromgevingen, onderwijs-op-maat, begeleiding in open afstandsonderwijs, numeriek rendement en studieloopbanen, strategische planning in het hoger onderwijs, leereffecten van stages en ervaringen met doceren door middel van een vreemde taal. Omdat het onmogelijk is alle presentaties te bespreken, beperk ik mij tot drie onderwerpen: onderwijs-op-maat, studierendement en kwaliteitszorg.

Onderwijs-op-maat is onderwijs dat optimaal is afgestemd op de studiebehoeften en -intenties van studenten. Door de Open Universiteit werd onderzocht of dit type onderwijs een zinvol antwoord is op de maatschappelijke behoefte aan flexibel onderwijs. Kenmerkend voor dit type onderwijs is dat studenten zelf een essentiële inbreng hebben in hun studieproces. Zij stellen niet alleen hun eigen leerdoelen vast, maar hebben ook een actieve inbreng bij de keuze van de leerstof, werkvormen, begeleiding en toetsing. In overleg met de docent/begeleider wordt een studietraject-op-maat uitgestippeld dat neergelegd wordt in een studiecontract (Schlusmans & Van den Boom). In theorie lijkt dit een aantrekkelijk concept, dat in de praktijk echter op een aantal problemen stuit. Men kan zich bijvoorbeeld afvragen of onderwijs-op-maat niet teveel eisen aan de student stelt en of het mogelijk is een dergelijke flexibiliteit te realiseren. Een ander punt is de waarde van een diploma van onderwijs-op-maat. Uit de behoeftepeiling die Van den Boom en Van der Velden (OU) naar dit type onderwijs deden, blijkt dit een belangrijk punt van twijfel. Onderwijs-op-maat zal dan ook waarschijnlijk beperkt blijven tot een flexibel keuzeprogramma dat enigszins aangepast kan worden aan de student. Het door Koper (OU) ontwikkelde computerprogramma waarmee de student zelf kan bepalen welke (invulling van een) studietaak het beste bij zijn/haar individuele behoeften en mogelijkheden aansluit, is hierbij een praktisch hulpmiddel. De door Schlusmans en Van den Boom in een experiment gehanteerde strategie om tot een zinvol studie-contract te komen, lijkt mij met name voor assistenten-in-opleiding aanbevelenswaardig, ook (of misschien zelfs) als zij deels een verplicht opleidingsprogramma moeten volgen.

Studierendement is een onderwerp dat door de ontwikkelingen op beleidsniveau sterk aan actualiteitswaarde heeft gewonnen. Veel onderwijsinstellingen laten momenteel onderzoek doen naar de oorzaken van lage rendementen. Voorbeelden zijn de onderzoeken van Peters (Hogeschool Nijmegen) en Koopman (UvA), waarin getracht wordt inzicht te krijgen in belemmerende en bevorderende studiefactoren, zowel bij de student zelf als in het onderwijsprogramma en de begeleiding. Bax en Van Drunen (TUE) onderzochten de verbetering van het rendement na een andere toetsprogrammering en Felix en Voorthuis (UvA) richtten zich in hun onderzoek op migrantenstudenten. Deze studenten zijn sterk ondervertegenwoordigd op de universiteit, maar de groep die daar eenmaal is beland, lijkt nauwelijks problemen te ondervinden. Deze bevinding wordt ondersteund door de resultaten van een studievoortgangsonderzoek uitgevoerd door Admiraal en Boogaard (IVLOS), dat zij presenteerden binnen het thema 'Onderwijs en Samenleving'. Uit hun onderzoek kwam naar voren dat allochtone studenten vergeleken met autochtone studenten geen studievertraging oplopen. Het rendementsprobleem ligt bij deze groep dus al eerder in de schoolloopbaan.

Naast onderzoek op instellingsniveau, werd ook breder onderzoek op het gebied van studierendement gepresenteerd. Door het CSHOB en IOWO is in opdracht van het Ministerie van O&W onderzocht in hoeverre een landelijke rendementsrapportage haalbaar is (Bijleveld, Jurgens & Prins). De onderzoekers lieten zich positief uit over de noodzaak van onderwijsin-

stellingen om open kaart te spelen over hun rendement. Voorwaarde is dan wel dat een dergelijk systeem betrouwbare gegevens oplevert. Dat is echter nog niet zo eenvoudig. De administratieve hindernissen lijken bepaald niet gering. Zo kwam uit een onderzoek van Wilbrink en De Jong (UvA) naar voren dat het verschijnsel 'no-show', studenten die ingeschreven staan bij een studie die zij niet meer volgen, met name een administratief probleem is. Studenten worden bij een wisseling van studie bijvoorbeeld pas aan het eind van het jaar uitgeschreven. In een reactie op Bijleveld et al. vroeg Wilbrink zich af op welke vragen een jaarlijkse rendementsrapportage antwoord zou moeten geven. En wat voor beslissingen zouden er aan moeten worden gekoppeld? Vossensteyn, Jongbloed en Koelman (CSHOB) pleitten voor een koppeling van de middelen die een onderwijsinstelling krijgt aan de studieduur van studenten. Het zou de instellingen er toe dwingen meer aandacht te besteden aan de studeerbaarheid van de aangeboden onderwijsprogramma's. Dat er ook een andere kant aan deze medaille zit bleek echter toen de discussiant, de heer De Vijlder van het Ministerie van O&W, aan het woord kwam. Om het toekennen van financiële middelen aan studierendement te kunnen koppelen, zouden onderwijsinstellingen de wettelijke mogelijkheid moeten krijgen studenten te weigeren op grond van vermeende ongeschiktheid, zo meende hij. Hoe dit precies plaats moet vinden is echter onduidelijk. Dat een dergelijke selectie positief uitpakt voor het studierendement wordt door onderzoek bovendien niet ondersteund. Daarentegen is de kans dat bepaalde groepen studenten (bijvoorbeeld allochtonen) hierdoor extra benadeeld worden niet denkbeeldig.

Tot slot wil ik nog enkele presentaties noemen op het gebied van *kwaliteitszorg*, eveneens een onderwerp dat al jaren meeloopt, maar nu opnieuw actueel is. Poell en Voeten (KUN) presenteerden een instrument om leereffecten van stages te meten en Wolfhagen en Van der Vleuten (RL) lieten zien bij welke aantallen studenten resultaten van evaluatie-onderzoek naar de kwaliteit van stages nog generaliseerbaar is. Weusthof, Frederiks en Westerheijden (CSHOB) keken naar de invloed van visitaties op de aandacht die onderwijsinstellingen aan kwaliteitszorg besteden. Zij concludeerden dat deze er met name voor zorgen dat hier meer aandacht voor komt, hoewel dat niet meteen geconcretiseerd wordt in maatregelen. Het stelsel van kwaliteitszorg is volgens hen dan ook 'meer dan een rituele dans'. Laten we hopen dat dat waar is, en dat kwaliteitszorg meer betekent dan simpelweg rendementsverbetering. Anders zou de toekomst er voor sommige groepen potentiële studenten wel eens somber uit kunnen zien.

Leren & instructie (J. van der Linden, Universiteit Utrecht)

Het is vrijdagmorgen 28 mei 1993, zo tegen 12.00 uur in zaal Maas/Rhein van het MECC in Maastricht. In de paneldiscussie over de bruikbaarheid van onderwijspsychologisch onderzoek voor de lerarenopleiding zit vaart. Het publiek krijgt de gelegenheid zich in de discussie te mengen. Boekaerts (RUL) brengt een persoonlijke ervaring in. "Ik kwam laatst op een feestje in gesprek met mijn nichtje. Wij spraken over hoe het met haar op school ging en over al het huiswerk dat ze voor school moest doen. Zij vertelde mij dat ze moeite had om met het huiswerk te beginnen. Als ze eenmaal begonnen was, dan had ze meestal geen moeite het ook af te maken. Ik vroeg haar of ze wist hoe dit zo kwam. Ze antwoordde dat ze dacht van wel. Ze had namelijk van school het advies gekregen om steeds met het moeilijkste vak te beginnen. Dan wist ze tenminste zeker dat ze dat af kreeg. Haar ervaring was echter dat ze tegen zo'n lastig begin zo opzag dat ze het uitstelde. Vaak tot aan het begin van de avond. Ik vroeg haar of ze een idee had wat voor haar beter zou kunnen werken bij de aanpak van het huiswerk. Ze zei dat ze dacht van wel. Ze kon wellicht beter beginnen met een eenvoudiger vak. Ze kwam er dan goed in en waarschijnlijk zou ze dan ook het moeilijkste vak beter aankunnen. Bovendien zou het voordeel zijn dat ze niet zo laat op de avond pas klaar zou zijn met het huiswerk. Ik heb haar aangeraden het voortaan ook maar zo te doen als ze nu zei".

Wat kan dit voorbeeld ons duidelijk maken over de verhouding tussen leren en instructie?

Op de eerste plaats dat het noodzakelijk is om na te gaan of een aanwijzing om op een bepaalde manier te leren ook leidt tot de gewenste leeractiviteiten en leerresultaten. Veelal wordt

te gemakkelijk volstaan met het opdragen van school- en huiswerk en het overhoren ervan. In veel gevallen heeft een docent geen weet van de manier waarop het (huis)werk wordt aangepakt, welke problemen zich voordoen bij het zich eigen maken van de inhoud, welke hulp noodzakelijk zou zijn. De uitkomst telt. Het proces blijft verborgen. Dit kan en mag niet zo zijn. De instructie dient *procesgericht* te zijn. Het gaat om maatregelen die aanzetten tot de gewenste leerprocessen. In nagenoeg alle bijeenkomsten binnen het thema 'Leren en Instructie' klonk deze boodschap door. Voor de goede orde, het ging om vijf thematische papersessies (lezen en tekstverwerking, rekenen en rekenvaardigheid, leer- en controle strategieën I en II, motivatie en inzet), een gestructureerde poster-sessie (domeinspecifieke onderwijsleerprocessen), 3 symposia (het verwerven van expertise in complexe domeinen, onderzoek naar literatuuronderwijs en Logo) en een keynotelezing (J.J. Elshout, UvA). In deze lezing met als titel *'Intentie en autonomie in het leerproces'* (verwijzend naar de noemer waaronder de bekende Utrechtse onderwijsleerpsycholoog C.F.M. van Parreren zijn proefschrift schreef) wees Elshout op de beperkingen van de situativiteitstheorie. Een theorie die thans, in het bijzonder in de Verenigde Staten, richtinggevend is voor onderwijs- en instructiepsychologen (zie *Cognitive Science*, 1993, volume *17*(1) voor een kritisch debat). Het is interessant te bemerken dat Elshout stelling neemt tegen een Amerikaans modernisme. Dat deed Van Parreren indertijd immers ook. De inhoud verschilt evenwel. Elshout vindt zijn inspiratie bij de informatieverwerkingsprocessen. Van Parreren had indertijd daar zijn twijfels bij en achtte het begrip handeling geschikter om zowel bedoelde (intentionele) als bepaalde (autonome) processen die zich bij het leren voltrekken te kunnen verklaren.

De belangstelling voor al deze sessies was over het geheel genomen heel behoorlijk. De zaal was vrijwel steeds goed gevuld. Het was ook verheugend te zien dat het leeuwedeel van de presentaties verzorgd werd door een junioronderzoeker met de senior op de achtergrond. Nagenoeg al deze presentaties waren 'to the point' en binnen de gestelde tijd. Het was duidelijk dat aan de presentaties van te voren zorg was besteed. Een prima zaak mijns inziens. Bovendien is het voorbeeldig als je spreekt over leren en instructie.

Wat toont het voorbeeld ons nog meer? Op de tweede plaats dat het klaarkomen met school- en huiswerk om een verstandige, met overleg gekozen, aanpak vraagt. Een aanpak die tussen leerlingen kan verschillen, mede afhankelijk bijvoorbeeld van het doel dat wordt nagestreefd (een hoog cijfer halen? of persoonlijke interesse?). De aanpak kan ook verschillen per vak of onderdeel binnen een vak. Afhankelijk bijvoorbeeld van hoe sterk aangesloten kan worden bij voorkennis. Bovendien is het gunstig als een leerling weet heeft van de aanpak of strategie die voor hem of haar met een grote waarschijnlijkheid tot het gewenste doel leidt. Zorg er dus in het onderwijs voor dat leerlingen handige *cognitieve strategieën* verwerven voor het bemeesteren van (probleem-)opgaven en daar ook weet van hebben. Deze (tweede) boodschap klonk in verschillende toonaarden door in de sessies; het meest pregnant in de sessie 'Leer-en controlestrategieën', meer domeinspecifiek gekleurd in de sessies 'Lezen en tekstverwerving', 'Rekenen en rekenvaardigheden' en 'Domeinspecifieke onderwijsleerprocessen'.

Het voorbeeld maakt nog meer duidelijk. Op de derde plaats dat het niet altijd eenvoudig is met (intentioneel) leren een begin te maken en het ook vol te houden. Daar komen emoties bij kijken, daar doen zich spanningen bij voor, spanningen die gereguleerd moeten worden om zich volledig te kunnen concentreren op het zich eigen kunnen en willen maken van de taakinhoud. Het onderwijs kan er niet aan voorbijgaan dat leren *de gehele persoon* betreft en dat een begeleiding en ondersteuning ook die gehele persoon moet dienen en niet alleen zijn of haar verstand. Deze (derde) boodschap is vooral tot uitdrukking gebracht in de sessie 'Motivatie en inzet'.

Ten slotte maakt het voorbeeld duidelijk dat een dialoog als instrument kan dienen om kennis te verwerven. Dit kan een dialoog zijn tussen een deskundige en een (goede) beginner, zoals in het voorbeeld. Het kan ook een dialoog zijn tussen personen die in kennis en vaardigheden op het betreffende domein niet zo veel van elkaar verschillen: leerlingen onderling bijvoorbeeld. Een vorm waarin zo'n dialoog tussen leerlingen onderling gestalte

kan krijgen is bijvoorbeeld het in afwisseling vervullen van een rol in activiteiten die gewenst zijn voor het begrijpen van een studietekst, zoals vragensteller en samenvatter in eigen woorden ('reciprocal teaching'). Het blijkt dat *sociale interactie* het leren en de ontwikkeling kan bevorderen. Er schuilt meer leer- en ontwikkelingspotentieel in interacties tussen leerlingen bij het gezamenlijk werken aan problemen dan menigeen in het onderwijs denkt of wenst te geloven. Benut derhalve dit potentieel zou een (vierde) boodschap kunnen zijn bij leren en instructie. Wij hebben deze boodschap daar dit jaar wat minder gehoord. Daarvoor moesten wij zijn bij 'de Maastrichtse inbreng' op deze ORD: het probleemgestuurd onderwijs (een mooie noemer overigens voor het werk van onze nieuwe staatsecretaris uit Maastricht voor Onderwijs en Wetenschappen). Ik zou evenwel buiten de grenzen van mijn bijdrage aan deze kroniek treden door hierover te rapporteren.

Methodologie & Evaluatie (P. Sanders, CITO, Arnhem)

Een groot aantal papersessies in deze divisie had betrekking op theoretische en praktische aspecten van studietoetsen en andere meetinstrumenten.

In de papersessie *'itemresponsmodellen'* vergeleek De Gruijter (RUL) parametrische (o.a. het Raschmodel) en het niet-parametrische itemresponsmodel met dubbele monotonie van Mokken. Indien genoemde modellen voldoen hebben parametrische modellen de voorkeur. Ook de lezing van Hemker (UU) ging over het model van Mokken, met name een uitbreiding van het model voor het analyseren van polytome itemscores. In de lezing werd uitgebreid ingegaan op toepassingen met behulp van computerprogrammatuur die momenteel beschikbaar is. Ook Tan en Imbos (RL) maakten bij hun onderzoek gebruik van een niet-parametrisch longitudinaal itemresponsmodel. Zij pasten dit model toe in het kader van het toetsingssysteem (voortgangstoetsen) zoals dat door de Rijksuniversiteit Limburg voor het meten van (kennis)groei gehanteerd wordt. Itemresponsmodellen was eveneens het onderwerp van de *key-note* lezing van Mellenbergh (UvA) die een algemeen conceptueel kader voor itemresponsmodellen presenteerde. Vele bekende psychometrische modellen blijken speciale gevallen van dit algemene kader te zijn. Bekende psychometrische begrippen zoals betrouwbaarheid en vraagonzuiverheid kunnen ook in dit kader ondergebracht worden. In de papersessie *'beoordelen'* deden Van den Bergh en Kuhlemeier (UU/CITO) verslag van een onderzoek naar het gebruik van rapportcijfers in onderwijsonderzoek, in dit geval cijfers voor aardrijkskunde. De relatie tussen toetsprestaties en rapportcijfers werd geanalyseerd met behulp van een multiniveau model met drie niveaus: leerling, klas en school. De conclusie was dat rapportcijfers gebrekkige indicaties zijn voor de rangordening van het prestatieniveau van leerlingen uit verschillende klassen of scholen. Klerks en Frijns (TUD) deden onderzoek naar beoordelingsmethoden bij het beoordelen van ontwerpopdrachten zoals die voorkomen bij de opleiding bouwkunde van de Technische Universiteit Delft.

'Toetsautomatisering in theorie en praktijk' was de titel van het symposium dat gewijd was aan het verschijnsel toetsservicesysteem. De belangstelling voor deze systemen neemt de laatste jaren toe in verband met de toenemende modularisering van verschillende onderwijsvormen. In verschillende lezingen kwamen zaken als itembanking en adaptief toetsen aan de orde. Een (Engelstalig) symposium ging over het gebruik van scoringsregels bij meerkeuzevragen dat als voordeel heeft dat leerlingen gedetailleerde informatie krijgen over elk item. De scoringsregels die door Dirkzwager, Leclerq en Shuford (UT/Université de Liège/Applied Probability Technology, Dallas, Texas) werden besproken gaan terug op het befaamde artikel van laatstgenoemde dat reeds in 1966 in het tijdschrift *Psychometrika* verscheen.

In de papersessie *'evaluatie van onderwijs'* deed Smeets (ITS) verslag van een onderzoek naar een toepassing van telematica in afstandsonderwijs, zogenaamd tele-leren. De conclusie was dat tele-leren niet als een volwaardige vervanger voor de traditionele les beschouwd kan worden.

In de papersessie *'toetsen en examineren'* ging Groen (CITO) in op kwaliteitscriteria voor centrale examens in het voortgezet onder-

wijs. De criteria zijn uitgewerkt in dertig aanbevolen procedures, bijvoorbeeld het rapporteren over lokale betrouwbaarheid en vraagonzuiverheid. De lezing van Kremers en Sluijter (CITO) ging over de constructie van toetsen ter afsluiting van de basisvorming die in 1993/94 ingevoerd wordt. Het primaire doel van de toetsing is na te gaan wat iedere leerling van elk vak in de basisvorming opgestoken heeft.

Het valideren van plaatsingstoetsen voor het Nederlands was het onderwerp van de lezing van Sluijter en Van Roosmalen (CITO). De beslissing die met de toetsen genomen wordt betreft doorverwijzing van leerlingen naar HAVO en VWO. Er bleken verschillen te bestaan tussen toetsscores enerzijds en prestaties van leerlingen na plaatsing en oordelen van docenten Nederlands anderzijds.

Een interessante toepassing van de generaliseerbaarheidstheorie bij het construeren van toetsen voor gespreksvaardigheid werd gepresenteerd door Bögels en Van Mourik (RL). Zijn lieten zien dat het mogelijk is de afnametijd te verkorten zonder dat daardoor de kwaliteit van het instrument niet meer acceptabel is.

Beroeps- en bedrijfsopleidingen (J. van der Sanden, Katholieke Universiteit Brabant, Tilburg)

Het programma van het onderdeel 'Beroeps- en bedrijfsopleidingen' omvatte een key-note lezing, 3 symposia, 9 paperpresentaties en een forumdiscussie.

De *key-note* lezing, getiteld 'learning in organizations', werd verzorgd door Mandl van de Ludwig-Maximilians Universiteit van München. In zijn betoog ging hij in op het onderscheid tussen impliciet/incidenteel leren en expliciet/intentioneel leren. Voor arbeidsorganisaties is het van belang dat er sprake is van een leercultuur, waarin beide vormen van leren kunnen gedijen. Volgens Mandl kunnen expliciete leerprocessen worden bevorderd door gebruik te maken van ideeën ontleend aan 'cognitive apprenticeship' en 'anchored instruction'. Hier ligt zijns inziens een belangrijk aandachtspunt voor opleiders en onderzoekers.

Het eerste symposium, onder voorzitterschap van Nijhof (UT) had als titel *'Eindtermen in het beroepsonderwijs'*. Achtereenvolgens werd aandacht besteed aan de sturende invloed die uitgaat van eindtermen bij enkele curriculumontwikkelingsprojecten (Frietman, ITS), de ontwikkeling van cascotoetsen op basis van eindtermen in het land- en tuinbouwonderwijs (Treep, CITO) en de wijze waarop bij enkele curriculumontwikkelingsprojecten in het technisch beroepsonderwijs eindtermen zijn vertaald in lesmateriaal (Nijhof). Uit het door Nijhof gerapporteerde onderzoek bleek onder andere dat niveau, aard en inhoud van in samenwerking met het bedrijfsleven gelegitimeerde eindtermen ingrijpend veranderden gedurende het curriculumontwikkelingsproces. Zo was er in het uiteindelijke lesmateriaal (vrijwel) niets terug te vinden van in eerdere stadia van het ontwikkelingsproces wenselijk geachte, op transfer gerichte doelstellingen.

Het tweede symposium (voorzitter Van der Sanden, KUB) had als thema *'Leren op de werkplek'*. In dit symposium werd empirisch onderzoek gepresenteerd naar de wijze waarop in een aantal grote arbeidsorganisaties gebruik wordt gemaakt van werkplekopleidingen. De Jong, Vendrig en Versloot (UU) stelden een classificatieschema op en probeerden hiermee varianten van werkplekopleidingen te beschrijven. Onstenk en Voncken (SCO) deden verslag van case studies, waarin werd nagegaan in welke mate de inrichting van de werkomgeving leerprocessen mogelijk maakt. Van der Sanden en Hornman (KUB) presenteerden de resultaten van een experimenteel onderzoek naar de effecten van de Leittextmethode.

In het derde symposium, getiteld *'Innovatieve scholingspraktijken in het bedrijfsleven'*, is verslag gedaan van een drietal door het ITS te Nijmegen uitgevoerde onderzoeken naar de aanpak van scholingsvraagstukken in het bedrijfsleven. Centraal stond de wijze waarop in de detailhandel (Van den Tillaart), de voedingsmiddelenindustrie (Feijen) en de autoreparatie- en distributiesector (Warmerdam) is gereageerd op technologische, organisatorische en/of marktontwikkelingen.

In de gepresenteerde *papers* kwamen uiteenlopende onderwerpen aan de orde: opleidingsbereidheid en -deelname in de financiële sector (Van Meensel, Hoger Instituut voor Arbeid, Leuven); de kwaliteit en beschikbaarheid van praktijkleerplaatsen en stageplaatsen voor leerlingen van het MBO (Moerkamp, UVA); de toepassing van het COMBI-model ten aan-

zien van sociaal-normatieve vaardigheden in het MBO (Verijdt, ITS); de toepassing van de Delphimethode bij onderzoek naar sleutelkwalificaties in het MBO (Van Zolingen & Klaassen, KUN). De overige papers hadden alle te maken met het thema 'evaluatie'. Door Bergenhenegouwen en Voorthuis (UVA), Bosman en Walpot (CITO) en Gelderblom en De Koning (NEI) werd ingegaan op de opzet en resultaten van verschillende bij bedrijfsopleidingen toegepaste evaluatiesystemen. Van Sandick en Schaap-Neuteboom (Albert Heijn) bespraken een methode voor het berekenen van het rendement van een training voor afdelingsmanagers. Wognum (UT) presenteerde de resultaten van een onderzoek naar het opleidingsbeleid en de effectiviteitsbepaling van een aantal opleidingsafdelingen van arbeidsorganisaties.

In de *forumdiscussie* debatteerden Van der Krogt (KUN), Nieuwenhuizen (EUR), Overduin (Vergouwen Overduin), Pieters (UT) en Thijssen (Rabo-bank) onder leiding van Simons (KUN) over het te voeren onderzoeksbeleid op het gebied van bedrijfsopleidingen. Enkele onderzoeksvragen die naar voren kwamen waren: Hoe verlopen autonome leerprocessen? Hoe kan het leervermogen van met name lager opgeleiden worden vergroot? Welke factoren bepalen de mate waarin transfer optreedt? Hoe kan de relatie tussen het beroepsonderwijs en arbeidsorganisaties worden verbeterd?

Probleemgestuurd Onderwijs (J. Willems, IOWO, Katholieke Universiteit Nijmegen)

In 1992 is tijdens de internationale ORD (ECER) in Enschede binnen het thema 'Hoger Onderwijs en Volwasseneneducatie' een symposium over probleemgestuurd onderwijs (PGO) gehouden. Dit symposium trok veel belangstelling van zowel presentatoren als deelnemers. Vanwege deze belangstelling en omdat de ORD'93 in Maastricht, de Nederlandse bakermat van het PGO, werd gehouden, vormde PGO op de ORD een eigen thema. De juistheid van deze beslissing bleek uit het aantal deelnemers: vrijwel voortdurend schommelde het aantal bezoekers rond de 50. Zowel in het WO als in het HBO blijkt PGO in toenemende mate ingevoerd te worden. Onder de deelnemers aan het thema waren dan ook zowel onderzoekers als praktijkmensen uit het HBO en WO en mensen die vooral geïnteresseerd waren in de mogelijkheden en moeilijkheden van PGO om daarop hun beslissing over de invoering van PGO te kunnen afstemmen. Al deze belangstellenden te zamen gaven ook vorm aan hetgeen de organisatoren met de ORD beoogden, namelijk een markt te zijn waarop onderzoekers, gebruikers en belangstellenden elkaar konden ontmoeten. Deze markt leerde overigens ook dat er vrijwel evenveel PGO-vormen zijn als er plaatsen zijn waar met PGO wordt gewerkt.

Tijdens het thema kwamen verschillende presentatievormen aan bod: de meer traditionele papers, een symposium rond een specifiek thema, in casu de voortgangstoets en verschillende ronde-tafel conferenties, waarin met een beperkt aantal deelnemers een concreet probleem uit het PGO werd besproken.

De gepresenteerde papers brachten vaak nieuwe resultaten. Dat lijkt misschien enigszins triviaal, maar een aantal jaren geleden had ik regelmatig de indruk dat op de ORD alleen presentaties aan de orde kwamen van al eerder gepubliceerde artikelen. Dat was nu dus duidelijk anders. Doorgaans was er dan ook sprake van een levendige discussie over mogelijke verklaringen en implicaties van de gepresenteerde resultaten. Met name de presentatie van Gijselaers en Schmidt (RL) over *tijdsbesteding en studieresultaten* had aanzienlijk meer discussietijd verdiend. Naast papers over onderzoeksresultaten, was er in de helft van de gevallen sprake van beschrijvende dan wel evaluatieve presentaties. Opvallend was dat in die gevallen de discussie wat vlakker leek dan bij de onderzoekspresentaties.

Bij het symposium over *'Voortgangstoetsen'* was het jammer dat er niet meer onderzoekers aanwezig waren. Alleen al uit onderzoeksoogpunt lijkt het alleszins aantrekkelijk te pleiten voor invoering van voortgangstoetsen in het onderwijs. Je krijgt over een periode van een aantal jaren de beschikking over vergelijkbare studieresultaten per student. Dit heeft bijvoorbeeld voordelen als je de resultaten van veranderingen in het onderwijsprogramma wilt vaststellen (voor/nameting; time-series). Overigens bleken natuurlijk ook de nodige haken en ogen te zitten aan voortgangstoetsen, bijvoorbeeld in opleidingen met (veel) differentiaties.

De *'ronde-tafel conferentie'* bleek een aardige vorm om van gedachten te wisselen over ervaren problemen en al dan niet succesvol gebleken oplossingen. Deze vorm werd door de deelnemers zo op prijs gesteld dat spontaan het voornemen ontstond dergelijke bijeenkomsten vaker, eventueel buiten het kader van de ORD te beleggen. Wel lijkt het dan verstandig te zorgen voor meer homogene deelnemersgroepen.

Kortom, het thema PGO bood voor de deelnemers een interessant en gevarieerd programma. Het is gezien het toenemende gebruik van PGO in Nederland in de toekomst vaker op de ORD te verwachten, zij het misschien niet als thema maar als uitgebreid symposium.

Aan deze kroniek werkten mee: *J. J. H. van den Akker, R. Bosker, G. ten Dam, J. van der Linden, G. Rijlaarsdam, P. Sanders, J. van der Sanden, P. J. J. Stijnen, J. Terwel, M. G. Weide, J. Willems*

De eindredactie werd verzorgd door *P. J. J. Stijnen*

Literatuur

Berkel, H. van (samensteller) (1993), *Onderwijsonderzoek in Nederland en Vlaanderen 1993. Proceedings van de Onderwijsresearchdagen 1993.* Houten/Zaventem: Bohn Stafleu Van Loghum.

Mededelingen

Inhoud andere tijdschriften

Comenius
13e jaargang, nr. 2, 1993

Thema: Alice is trying something

Emily, Sarah and Alice: models for girlhood and education. An introduction, door M. van Essen, M. Volman en M. Lunenberg
The modern schoolgirl. The schooling of young women in Australia, door L. Johnson
'Just as good as the males but without their competition'. A pluralist look at female achievement, door L. A. Bell
Class and gender in the classroom, door H. Bjerrum Nielsen en M. Rudberg
Gender and prosocial development in different educational settings, door E. Öhrn
Girls in the educational research discourse, door M. Volman, G. ten Dam en E. van Eck
De speurtocht naar volwassenheid. Het volwassenheidsconcept binnen het pedagogisch vertoog, 1945-1970, door Y. te Poel en S. Miedema

Pedagogisch Tijdschrift
18e jaargang, nr. 2, 1993

Opvoedingsvoorlichting. Op zoek naar theoretisch bruikbare kaders en modellen, door J. H. A. Groenendaal
Over de instructieve waarde van opvoedingsvoorlichting voor ouders, voorlichters en onderzoekers. Evaluatie van een project, door M. W. E. Lambermon
Kritische beschouwing van een aantal projecten m.b.t. opvoedingsvoorlichting, door S. Roomans
Opvoedingsvoorlichting in de praktijk. Van opvoedingsvoorlichting naar opvoedingsondersteuning, door I. Bakker
Onzekere ouders? Over de geschiedenis van opvoedingsvoorlichting en opvoedingsonzekerheid, door N. Bakker

Pedagogisch Tijdschrift
18e jaargang, nr. 3, 1993

Pedagogisch commentaar, door A. Snick
Fouten bij het oplossen van elementaire algebra problemen en de invloed van groepswerk, door J. Chr. Perrenet
Doordenken over deregulering, door P. Sleegers en A. Wesselingh
Onderwijsplanning: zelfreferentiële problemen, door R. Vanderstraeten

Tijdschrift voor Onderwijsresearch
18e jaargang, nr. 2, 1993

Themanummer Opportunity structures in primary schools

Improving opportunities for the educationally disadvantaged: introduction to this special issue, door M. P. C. van der Werf
Effective provisions for ethnic minorities in primary schools, door M. G. Weide
The effect of teachers' attitudes and practices on the achievements of Dutch low SES pupils, door L. F. J. van der Velden
Pygmalion and the effectiveness of 'black' schools; teachers' sterotypes and hidden goal differentiations towards ethnic minorities, door P. Jungbluth
Secondary schoolrecommendations in relation to student and school caracteristics, door L. Mulder
Cooperation between Educational Priority Policy areas and basic adult education, door G. J. Reezigt

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 7/8, 1993

Jongeren aan de rand van de samenleving, door O. van Heffen
Evaluatie van een zorgverbredingsprogramma voor basisschoolteams, door A. Jager en L. M. Stevens
Basisdiagnostiek in de psychosociale jeugdhulpverlening II: gebruik van de VSPS in de Gezinsvoogdij, door E. M. Scholte

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 9, 1993

Leerlingvolgsystemen: een antwoord op schoolevaluatie? door A. B. van Doorn en F. J. H. Harinck
Risico-lezers: naar een succesvolle aanpak, door N. Osinga
Jeugdgezondheidszorg: doelstelling, organisatie en werkwijze, door H. Talma en R. A. Hirasing

Swets & Zeitlinger BV
Postbus 825 •
2160 SZ Lisse •
tel. 02521-35111 •
fax. 02521-15888 •



Gedragstherapie in de klas

H. van Bilsen, T. Engelen, S. Volker (redactie)

Gedragstherapeutische kennis en vaardigheden worden veelal gebruikt voor het aanpakken van individuele problemen. In veel mindere mate worden zij toegepast in het onderwijs. Toch zijn er voor het leertheoretisch referentiekader dat wel wordt aangeduid met de term Applied Behavioral Analysis (ABA) verschillende toepassingsmogelijkheden in het onderwijs. De verschillende bijdragen getuigen hiervan. Vanuit deze gedragstherapeutische invalshoek komen aan de orde de oplossing van problemen in het voortgezet onderwijs, de samenwerking tussen school en gezin, het planmatig invoeren van leertheoretische interventies op basisscholen, begeleidingsinstrumenten vanuit interactioneel perspectief, behandeling van leerproblemen en de samenhang met gedragsproblemen, een interventieprogramma voor tekorten in sociale vaardigheden, een diagnostisch instrument voor en de behandeling van faalangstproblemen, een overzicht van de effecten van gedragsmodificatie in het onderwijs en een overzicht van de resultaten van twintig jaar ABA in het onderwijs.

f 45,00, 210 pag., ISBN 90 265 1300 3

Opvoeding en opvoedingsproblemen in het gezinsleven

P.M. Schoorl

Kinderen opvoeden is nauw verweven met het gezinsleven. Meestal vraagt men zich pas af hoe het opvoedingsproces verloopt als er problemen ontstaan. Dit boek geeft echter geen adviezen over de beste manier om concrete opvoedingsproblemen aan te pakken. Het biedt een theoretisch kader om de invloed van het gezin op de ontwikkeling van kinderen tot mondige mensen te onderzoeken. Centraal in het pedagogisch gezinsonderzoek staat het kritisch kijken naar de regels in het gezinsleven. Juist door de regels en de sociale orde van het gezin te onderzoeken, wordt zichtbaar hoe opvoeders ermee om kunnen gaan. Op die basis worden suggesties gegeven aan hulpverleners voor het analyseren van de specifieke gezinssituaties waarmee ze te maken krijgen.

f 45,00, 184 pag., ISBN 90 265 1265 1

Verkrijgbaar in de boekhandel of bel: 02521-35111

Swets & Zeitlinger *Publishers*
Postbus 825, 2160 SZ Lisse, tel. 02521-35111, Fax 02521-15888

Ts oct 6441

Inhoudsopgave

Pedagogische Studiën Jaargang 70, nummer 6/1993

R. Yap en A. van der Leij
Computergestuurde remediëring van dyslexie door het opvoeren van de herkenningssnelheid van subwoordeenheden 402

S. E. Sliepen en P. Reitsma
Training van lom-leerkrachten in directe instructie van begrijpend leesstrategieën 420

H. Dekkers
Determinanten van seksespecifieke vakkenkeuzen in HAVO en VWO 445

Kroniek
N. Verloop (red.)
Het AERA-Congres 1993, Atlanta 460

Boekbespreking 475

Mededelingen 478

Computergestuurde remediëring van dyslexie door het opvoeren van de herkenningssnelheid van subwoordeenheden

R. Yap en A. van der Leij

Samenvatting

In twee experimenten wordt nagegaan of een computergestuurde subwoordtraining leidt tot een verbetering van processen die ten grondslag liggen aan snel en efficiënt lezen. De training is gegeven aan dyslectische lezers die gemiddeld 10 jaar oud zijn en een leesachterstand van minstens twee jaar hebben. In de training wordt de herkenningssnelheid opgevoerd van subwoordeenheden die van orthografische of fonologische aard zijn, of een combinatie van beide. Het onderzoek laat zien dat de training leidt tot een snellere herkenning van individuele letters, subwoordeenheden, woorden en pseudowoorden. Wanneer zowel orthografische als fonologische subwoordeenheden als de koppeling daartussen getraind worden is er zelfs een surplus ten opzichte van een controletraining waarin eveneens intensief met woorden geoefend wordt op de computer. De subwoordtraining heeft vooral effect op het lezen van bestaande woorden die geïsoleerd worden aangeboden.

Inleiding

Dyslectische kinderen kampen met ernstige lees- en spellingsproblemen. Naast het maken van veel fouten blijken ze vooral problemen te ondervinden wanneer woorden snel en efficiënt gelezen moeten worden. Zij blijven vaak letter-voor-letter lezen. Accuraat respectievelijk snel lezen worden gezien als twee opeenvolgende fasen in de ontwikkeling naar het geautomatiseerd lezen (Ehri & Wilce, 1983; LaBerge & Samuels, 1974). Een lage accuratesse en/of snelheid van woordverwerking zijn aanwijzingen dat de basale processen van woordherkenning niet geautomatiseerd zijn. Onder basale processen van woordherkenning wordt verstaan de verwerking van de orthografische structuur van de taal (de volgorde van letters), de fonologische structuur (de volgorde van klanken) en de koppeling tussen orthografische en fonologische structuur (het decoderen).

De langzame uitvoering van de basale processen bij dyslectische kinderen is aangetoond in onderzoeken waaruit blijkt dat deze kinderen traag zijn in het lezen van woorden en pseudowoorden (Bouma & Legein, 1980; Lovett, 1987; Manis, 1985; Yap & Van der Leij, 1993) en het benoemen van niet-alfabetisch materiaal zoals cijfers, letters, plaatjes, en kleuren (e.g. Bowers & Swanson, 1991; Denckla & Rudel, 1976; Spring & Davis, 1988; Wolf, Bally & Morris, 1986). De trage leessnelheid van woorden geeft aan dat dyslectische kinderen problemen hebben met het snel verwerken van de orthografische structuur van woorden. De langzame verwerking van pseudowoorden is een indicatie dat het proces van decoderen traag verloopt bij deze kinderen. Decoderen is het toepassen van grafeem-foneemkennis bij het identificeren van woorden. Daar pseudowoorden geen betekenis maar wel een klankzuivere spellingstructuur hebben, kunnen ze alleen gelezen worden via het proces van grafeem-foneemkoppelingen. De lage benoemsnelheid van niet-alfabetisch materiaal geeft ten slotte aan dat ook de verwerkingstijd van puur fonologische processen, dus los van de koppeling met orthografie, lang is.

Naast een trage verwerkingssnelheid van de basale elementen van woordkennis is ook een inefficiënte codering van woorden een kenmerk van ongeautomatiseerd lezen (Baddeley, 1986; Frederiksen, 1982; LaBerge & Samuels, 1974). Het connectionistisch model biedt ruimte om beide kenmerken in hun onderlinge relatie te beschrijven en zal daarom in het kort besproken worden. Een efficiënte codering

houdt in dat woorden gelezen worden via clusters van letters, ofte wel subwoordeenheden, en niet op basis van afzonderlijke letters. Een inefficiënte codering leidt tot het letter-voor-letter lezen, een efficiënte codering via de detectie van subwoordeenheden tot directe woordherkenning. Volgens het model is de vorming van letterclusters gebaseerd op de herkenningstijd van de afzonderlijke letters (Adams, 1990). Stel dat een cluster uit twee letters bestaat. Wanneer de afzonderlijke letters snel herkend worden is de tijd tussen de activatie van beide letters in het geheugen netwerk kort en worden beide letters als het ware gelijktijdig, dus als één cluster, waargenomen. De sterkte van de associatie tussen de twee letters is verder afhankelijk van de frequentie waarmee de letters in combinatie met elkaar voorkomen in de schrijftaal. Hoe hoger die frequentie, hoe groter de kans op clustervorming is door herhaald lezen. Het model impliceert tevens dat orthografische clusters van twee en drie letters, de zgn. bi- en trigrammen, de basiselementen zijn waarop directe woordherkenning stoelt (Seidenberg & McClelland, 1989). Tot slot geeft het model aan dat er voortdurend associaties gelegd worden tussen de verschillende basale elementen van woordkennis: de orthografie, de fonologie, en de koppeling daartussen. Hierdoor is het mogelijk dat de versterking van één element leidt tot versterking van de andere elementen. Een verbetering in foneemherkenning kan bijvoorbeeld leiden tot een snellere letterherkenning en een verbetering van de decodeervaardigheid. De vraag is nu in hoeverre dit idee ook opgaat voor dyslectische kinderen, van wie aangenomen kan worden dat ze een fonologisch tekort vertonen (Snowling, 1987) en die wellicht minder beïnvloedbaar zijn via de fonologische weg.

Bovenstaande onderzoeken suggereren dat dyslectische lezers moeite hebben met het automatisch verwerken van woorden omdat ze subwoordeenheden niet snel kunnen herkennen. Deze subwoordeenheden zijn van orthografische aard (bv. het lettercluster *ze*) of fonologische aard (bv. het gesproken foneemcluster /om/) of een combinatie van beide (bv. de koppeling tussen het foneemcluster /om/ en het lettercluster *om*). Een training in het snel herkennen van dergelijke subwoordeenheden zou derhalve kunnen leiden tot een verbetering van de processen die ten grondslag liggen aan efficiënt lezen.

De hoofdvraagstelling van dit onderzoek luidt: leidt een training in het snel herkennen van subwoordeenheden tot een verbetering van processen die ten grondslag liggen aan efficiënt lezen? In het eerste experiment wordt in drie aparte, computergestuurde trainingen geoefend met subwoordeenheden die van orthografische of fonologische aard zijn of een combinatie van beide. In het tweede experiment worden daarnaast alle drie genoemde aspecten van woordkennis in één training gegeven. De onderliggende processen van efficiënt lezen worden geëvalueerd aan de hand van taken die verschillende vaardigheden meten, zoals de herkenning van individuele letters, subwoordeenheden, woorden en pseudowoorden. Op alle taken worden zowel accuratesse als snelheid gemeten. De experimenten staan ook beschreven in de dissertatie van de eerste auteur (Yap, 1993).

De keus voor computergestuurde instructie is gemaakt op grond van het feit dat daarmee een zo hoge beheersbaarheid van in- en output kan worden bereikt dat dit soort theoretisch relevante vraagstellingen beantwoordbaar worden. Daarnaast is uit ander onderzoek bij dyslectische leerlingen gebleken dat er met behulp van computergestuurde instructie inderdaad effect te boeken is. (vgl. o.a. Van den Bosch, 1991; Van Daal, 1993).

1 Experiment 1

1.1 De vraagstelling

In dit experiment worden drie varianten van een subwoordtraining vergeleken met een controletraining. In de subwoordtraining nemen kinderen deel aan één van de volgende experimentele trainingsvormen: a) een auditieve training waarin de fonologische structuur getraind wordt, b) een visuele training waarin de nadruk gelegd wordt op de orthografische structuur, en c) een auditief-visuele training waarin de letter-klankstructuur getraind wordt. De experimentele trainingen zijn afgeleid van de perceptual units training van Frederiksen, Warren en Roseberry (1985); Klapwijk en Das-Smaal (1990). In het kort komt het er op neer dat kinderen zo goed en zo snel mogelijk subwoord-

eenheden in woorden moeten identificeren. Het effect van de training wordt op twee manieren getoetst. Ten eerste worden de drie varianten van de subwoordtraining vergeleken met een controletraining die geen expliciet beroep doet op bovenstaande karakteristieken maar waarin eveneens intensief met woorden wordt geoefend op de computer (COPAL: Van Daal, Van der Leij, Bakker & Reitsma, 1987; Van der Leij, 1990; Smeets & Van der Leij, 1993). Door de vergelijking met de controletraining worden factoren uitgeschakeld als intensieve leesinstructie, het gewend zijn aan de computer en de hoeveelheid individuele aandacht. Ten tweede wordt het effect van de training getoetst door een retentiemeting die drie maanden na het stopzetten van de training plaatsvindt. Wanneer de afwezigheid van de experimentele training tot een kleinere vooruitgang in leesprestatie leidt dan de training zelf is het zeer waarschijnlijk dat de vooruitgang net na de training werkelijk toe te schrijven is aan specifieke effecten van de training.

Het experiment moet antwoord geven op de volgende vragen.

1. Leidt de experimentele training tot het beter herkennen van subwoordeenheden in zijn algemeenheid? Als dat zo is, zal de training niet alleen leiden tot een direct effect waarbij de getrainde subwoordeenheden beter herkend worden maar ook tot een indirect effect waarbij de training generaliseert naar subwoordeenheden die niet in de training gepresenteerd worden.
2. Is er een transfereffect naar het hardop lezen van woorden en naar de snelheid van letterherkenning? Een transfereffect betekent dat de training effect heeft op andere vaardigheden dan die welke rechtstreeks in de training worden geoefend. In de training wordt de herkenning van subwoordeenheden geoefend, niet het hardop lezen en ook niet het herkennen van individuele letters.
3. Passen de kinderen hun kennis van de geleerde subwoordeenheden toe bij het lezen? Als dat zo is, worden woorden die een getraind cluster bevatten beter gelezen dan woorden die een ongetraind cluster bevatten.

1.2 Methode

Proefpersonen

De proefpersonen zijn 78 dyslectische lezers uit het LOM-onderwijs die gemiddeld 10 jaar oud zijn en minstens twee jaar leesachterstand hebben. De proefpersonen zijn verdeeld over drie experimentele condities en een controletraining. De groepen zijn gematcht op leeftijd en technische leesvaardigheid. De technische leesvaardigheid is gemeten met de Eén-Minuut-Test (EMT; Brus & Voeten, 1973). De score op de EMT is uitgedrukt in het aantal goed gelezen woorden per minuut (wpm.). Uit een variantie-analyse met leeftijd en EMT-scores als afhankelijke variabelen is gebleken dat er geen significante verschillen zijn tussen de vier trainingsgroepen ($F < .10$). De matching op leeftijd en technische leesvaardigheid is dus geslaagd. De groepen verschilden ook niet significant op verbale intelligentie, gemeten met de Peabody Picture Vocabulary Test (PPVT: Dunn & Lloyd, 1965). Zie Tabel 1.

Algemene procedure

Het design is een pretest-training-posttest design met een retentiemeting. De volgorde van afnames is: voortoets (t1) - training - natoets (t2) - retentietoets (t3). De trainingsduur is 160 minuten, opgedeeld in 16 sessies van 10 minuten. De tijd tussen t1 en t2 bedraagt twee maanden. De tijd tussen de eindtoets en de retentietoets is drie maanden. In de retentietoets wordt de kinderen slechts gevraagd de leestaken uit te voeren (zie toetsen). De trainingen en de taken van de toetsen, behalve de Eén-Minuut-Test,

Tabel 1
Leerlinggegevens: gemiddelden per experimentele en controle-conditie

Training	*N*	Leeftijd (mnd.)	EMT (wpm.)	PPVT (IQ)
a	18	125.2 (9.57)	28.3 (10.5)	104.8 (9.25)
av	19	123.6 (9.32)	27.0 (9.10)	111.3 (10.7)
v	19	124.4 (9.99)	26.8 (10.0)	106.0 (6.11)
c	22	121.5 (9.46)	27.7 (7.34)	101.3 (11.7)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training en c = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

worden gepresenteerd via een Apple Macintosh computer. Gesproken woorden en auditieve subwoordeenheden krijgen de kinderen via een koptelefoon te horen. Deze spraakstimuli bestaan uit natuurlijke spraak die door middel van een analoog-digitaal converter is gesampled met een frequentie van 22 KHz. en opgeslagen op de harde schijf van de computer.

Trainingstaken

1. Auditieve training. Een kind krijgt eerst een foneemcluster te horen via een koptelefoon en 700 ms. later hoort hij een woord. Aan het kind wordt gevraagd om, voordat het een pieptoon hoort, aan te geven of de klank in het gesproken woord zit, bv. zit /om/ in /bekommeren/? Afhankelijk van de nauwkeurigheid van werken wordt de tijd tussen de aanbieding van het gesproken woord en het moment waarop het kind de pieptoon hoort steeds verkort.

2. Visuele training. Een kind krijgt op het scherm een lettercluster en daaronder een woord te zien. Aan het kind wordt gevraagd of het cluster in het woord zit. Afhankelijk van de nauwkeurigheid van werken wordt de aanbiedingstijd van het woord steeds verkort ('flitsen'). Het cluster blijft op het scherm staan.

3. Auditief-visuele training. Een kind krijgt eerst een foneemcluster te horen via een koptelefoon en 700 ms. later ziet hij een woord op het scherm. Aan het kind wordt gevraagd of de klank die hij hoort ook in het woord zit als je het woord zou uitspreken. Afhankelijk van de nauwkeurigheid van werken wordt de aanbiedingstijd van het woord steeds verkort ('flitsen').

4. Controletraining. Kinderen uit de controlegroep worden getraind met het eveneens computergestuurd trainingsprogramma COPAL. De training bevat allerlei leesoefeningen waarin stimuli auditief en visueel worden aangeboden: spellen, structuurrij-oefeningen, auditief-visueel matchen, etc. Het is, in tegenstelling tot de hiervoor beschreven trainingen, een multicomponentieel programma waarin diverse lees- en spellingvaardigheden worden getraind. Daarbij is zowel de aandacht voor snelheid als voor sub-woordeenheden geringer. Dit programma is effectief gebleken bij zwakke lezers (met een jaar achterstand) en in mindere mate bij dyslectici die vergelijkbaar waren met de leerlingen die in de onderhavige experimenten centraal staan (Smeets & Van der Leij, 1993). Het programma is, evenals de experimentele trainingen, volstrekt interactief: de leerling oefent zonder begeleiding of tussenkomst van een proefleider.

Materiaal en procedure van de training

In totaal worden er 10 verschillende clusters getraind; vijf van het type medeklinker-klinker en vijf van het type klinker-medeklinker. De visueel aangeboden clusters zijn: ag, an, ar, di, do, el, ke, or, vo, ze. Deze clusters worden in de visuele training gebruikt. De auditieve clusters zijn: aan, aar, ak, boo, is, naa, om, va, wa, we. Deze clusters worden in de auditieve en de auditief-visuele training gebruikt. Bij elk verschillend cluster hoort een apart databestand van 216 verschillende woorden. De woorden hebben één tot drie lettergrepen en worden één voor één at random gepresenteerd.

De training is een keer voorgedaan door de proefleider. Per dag vindt er per kind één sessie plaats. In één sessie worden vijf series afgewerkt. In een serie wordt één cluster gepresenteerd dat in 27 opdrachten, telkens bij verschillende woorden, gedetecteerd moet worden. Er zijn 'gelijke' opdrachten waarin een cluster inderdaad in een woord voorkomt en 'ongelijke' waarbij dat niet het geval is. In een serie zijn de 'gelijke' en 'ongelijke' opdrachten zoveel mogelijk evenredig verdeeld. In de 'gelijke' pogingen komen de te identificeren visuele clusters even vaak voor in de begin-, midden-, en eindlettergreep van een woord. De te identificeren auditieve clusters komen even vaak voor in lettergrepen waar de klemtoon op valt als in lettergrepen waar de klemtoon niet op valt, bijvoorbeeld /wa/ komt voor in /wasbak/ en in /afwassen/. In de 'ongelijke opdrachten' van de visuele training bevat de helft van de woorden een cluster dat lijkt op een cluster in het aangeboden woord, bijvoorbeeld het cluster *el* wordt gepresenteerd en daaronder het woord *leger*, waarin het afleidende cluster *le* zit. Na een serie wordt een volgend cluster aangeboden. Per sessie kwamen vijf van de tien experimentele clusters aan bod. De volgorde van aanbieding van clusters is steeds gerandomiseerd zonder teruglegging van het voorgaande cluster. Telkens als alle tien de clusters geweest zijn start de random-procedure opnieuw. Om de twee sessies zijn dus alle tien de clusters even vaak aan

bod gekomen. In een trainingsduur van 16 sessies is ieder cluster 8 keer aan bod gekomen. Het kind geeft antwoord door op een ja/nee knop te drukken.

Via een tellerstand en een afbeelding van drie, aanvankelijk zwart weergegeven, ballen krijgt het kind na iedere opdracht feedback over zijn snelheid en nauwkeurigheid van werken. De tellerstand is een indicatie voor de toegelaten antwoordtijd van de volgende opdracht, d.i. de tijd waarbinnen het kind antwoord moet geven. Deze antwoordtijd bedraagt 1/tellerstand × 60 sec. Een tellerstand van 20 komt dus overeen met een toegelaten antwoordtijd van 3 sec., d.i. 20 wpm. In de visuele en de auditief-visuele training betekent dit dat het woord 3 sec. op het scherm blijft staan. Als het kind na die 3 sec. nog antwoord geeft is het antwoord fout. In de auditieve training krijgt het kind na 3 sec. een toon te horen ter beëindiging van de antwoordtijd. Als een opdracht fout is beantwoord, daalt de tellerstand met 2 woorden per minuut (wpm.) en wordt één van de drie ballen grijs gemaakt. Bij een goed antwoord stijgt de tellerstand met 2 - de toegelaten antwoordtijd wordt dus korter - en wordt een grijze bal zwart gemaakt. Wanneer alle drie de ballen grijs zijn daalt de tellerstand met 20 en kan het kind vanuit een aanzienlijk lager niveau zijn snelheid weer proberen op te voeren. De minimale tellerstand is 10.

Toetsen

A. Detectietaken

Deze taken worden gegeven ter bepaling van het directe en indirecte trainingseffect. De computer registreert het aantal fouten en de latentietijden. Analoog aan de drie experimentele trainingsvarianten zijn er drie detectietaken. Het kind moet dus weer met behulp van een ja/nee knop aangeven of een cluster in een woord zit. Er worden zowel getrainde als ongetrainde clusters aangeboden. Elke variant bestaat uit 40 items.

1. Auditieve Detectietaak. Deze taak is analoog aan de auditieve training. Een kind hoort via een koptelefoon een klankcluster en na 700 ms. een woord. Deze taak wordt gegeven aan kinderen uit de auditieve en de controlegroep.

2. Visuele Detectietaak. Deze taak is analoog aan de visuele training. Een kind ziet een cluster en daaronder een woord. Zowel cluster als woord blijven op het scherm staan totdat het kind antwoord heeft gegeven. Deze taak wordt gegeven aan kinderen uit de visuele en de controlegroep.

3. Auditief-Visuele Detectietaak. Deze taak is analoog aan de auditief-visuele training. Een kind hoort een klankcluster en 700 ms. later ziet hij een woord op het scherm. Het woord blijft op het scherm staan totdat het kind antwoord heeft gegeven. Deze taak wordt gegeven aan kinderen uit de auditief-visuele en de controlegroep.

B. Transfertaken

Deze taken worden gegeven ter bepaling van transfereffecten naar leesvaardigheid (taak 1 t/m 3) en letterherkenning (taak 4). Taak 1 t/m 3 wordt ook gebruikt in de retentietoets.

1. Eén-Minuut-Test (EMT, vorm B; Brus & Voeten, 1973). Kinderen moeten zo goed en zo snel mogelijk woorden die onder elkaar staan in één minuut lezen. De score wordt uitgedrukt in het aantal correct gelezen woorden per minuut (wpm.).

2. Flitstaken. Er zijn twee flitstaken, van elk 20 items: a) 200 ms. Flitstaak Met Woorden, b) 200 ms. Flitstaak Met Pseudowoorden. De algemene procedure is dat woorden en pseudowoorden 200 ms. worden aangeboden op de computer. Daarna wordt de stimulus gemaskeerd door presentatie van een nonsense figuur die bestaat uit lijnen en cirkels, zodat het nabeeld op het netvlies verdwijnt. Het kind moet zeggen welk woord hij gezien heeft. De woorden en pseudowoorden bevatten zowel getrainde als ongetrainde clusters. De leesprestatie wordt uitgedrukt in het percentage goed gelezen woorden. Er wordt verondersteld dat het niet accuraat kunnen lezen van geflitste woorden te wijten is aan onvoldoende tijd om het woord te verwerken tot aan het moment waarop het maskeerteken verschijnt.

3. Benoemtaken. Er zijn twee benoemtaken, die elk 20 items bevatten: a) Woorden Benoemen, b) Pseudowoorden Benoemen. De algemene procedure is dat woorden en pseudowoorden voor onbepaalde tijd op een computerscherm blijven staan totdat het kind antwoord heeft gegeven. De woorden en pseudowoorden bevatten zowel getrainde als ongetrainde clusters. De nauwkeurigheid van de leesprestatie wordt uitgedrukt in percentage

goede antwoorden. De leessnelheid is de latentietijd van de goede antwoorden, uitgedrukt in milliseconden. De latentietijd is de tijd tussen de aanbieding van het woord en het moment waarop het kind het woord begint uit te spreken en is gemeten met behulp van een voicekey.

4. Letters Benoemen. Alle letters plus alle gecombineerde letters (lange klanken en tweetekenklanken), zoals 'ou', 'ij', 'aa', etc. worden één voor één op het scherm gepresenteerd. De letters blijven op het scherm staan totdat het kind antwoord heeft gegeven (36 items). De proefleider registreert het goede/foute antwoord op de computer. De latentietijd is gemeten met behulp van een voicekey en wordt direct geregistreerd.

Materiaal van de toetsen

De woorden op de detectie-, flits- en benoemtaken worden at random één voor één gepresenteerd op een computerscherm. De helft van de woorden bevat een getraind cluster, de andere helft een ongetraind cluster. De ongetrainde clusters die als equivalent van de getrainde auditieve clusters dienen zijn: ik, ien, iet, hu, ol, op, pa, po, raa, rui. De equivalenten van de getrainde visuele clusters zijn: ba, er, et, in, it, le, on, pi, to, wi. Elke taak start met 4 oefentrials. Het gebruikte lettertype is Amsterdam (kleine letters, puntgrootte 48).

Elke variant van de detectietaak bestaat uit 40 trials, d.i. de aanbieding van 40 clusters en 40 woorden. De woorden variëren in lengte van 3 tot 10 letters en zijn nog niet eerder gepresenteerd in de training. De 40 trials zijn evenredig verdeeld over 'gelijke' en 'ongelijke' opdrachten en over getrainde en ongetrainde clusters. De opbouw van 'gelijke' en 'ongelijke' opdrachten is analoog aan die van de training (zie materiaal en procedure van de training).

Het materiaal op de flitstaak en de benoemtaak bestaat per taak uit 20 woorden en 20 pseudowoorden die niet eerder gepresenteerd zijn in de training. Pseudowoorden zijn van het type mmkm (bijvoorbeeld 'flis') en mkmm (bijvoorbeeld 'pilm'). Bestaande woorden hebben een tweelettergrepige structuur waarbij iedere lettergreep een mkm- of een mkmm-structuur heeft, bijvoorbeeld 'tuinman', 'balkon', 'zandweg', etc. Verondersteld wordt dat een dergelijke structuur van woorden niet te makkelijk en niet te moeilijk is om enige variatie in scores te vinden bij de betreffende populatie zeer zwakke lezers.

1.3 Resultaten

Detectietaken

Op iedere detectietaak is een 2 (training) × 2 (tijdstip) ANOVA uitgevoerd op *A'* scores (Pollack & Norman, 1964) en latentietijden. De factor training is de tussen-proefpersoonvariable met twee niveaus: één van de drie experimentele trainingen vs. de controletraining. De factor tijdstip is de binnen-proefpersoonvariabele en heeft twee niveaus: voortoets vs. natoets. Tevens is bij de experimentele trainingsgroepen het beslissingscriterium berekend, uitgedrukt in B scores. Een hoog criterium (B > 1) geeft aan dat de proefpersoon voorzichtig is in het nemen van een beslissing. Een laag criterium (B < 1) geeft aan dat de proefpersoon veel risico neemt. In Tabel 2 zijn de scores weergegeven die behaald zijn bij het uitvoeren van de detectietaken.

1. Auditieve Detectietaak. Op de *A'* scores zijn geen significante effecten gevonden. Op de latentietijden zijn alle hoofdeffecten significant en is er een significante training × tijdstip interactie ($F(1,38) = 4.57, p < .05$). De interactie toont aan dat de auditieve trainingsgroep na de training sneller gesproken klanken in woorden kan identificeren dan de controlegroep. Het beslissingscriterium is daarbij gelijk gebleven.

2. Visuele Detectietaak. Er is een marginaal significante training × tijdstip interactie op zowel de *A'* scores ($F(1,39) = 4.06, p = .06$) als op de latentietijden ($F(1,39) = 27.1, p < .001$). Uit de interacties blijkt dat de visuele trainingsgroep na de training meer fouten maakt maar een lagere latentietijd heeft, en dus sneller presteert, dan de controlegroep. Ook daalt het beslissingscriterium significant bij de visuele groep ($F(1,39) = 8.53, p < .01$).

3. Auditief-Visuele Detectietaak. Er is een significante training × tijdstip interactie op zowel de *A'* scores ($F(1,39) = 4.70, p < .05$). als op de latentietijden ($F(1,39) = 16.0, p < .001$). Uit de interacties blijkt dat de auditief-visuele trainingsgroep na de training meer fouten maakt maar sneller presteert dan de controlegroep. Ook daalt het beslissingscriterium significant bij de auditief-visuele groep ($F(1,39) = 7.23, p < .01$).

Tabel 2
Scores op de Detectietaken

Taak	Training	A'		Latentietijd	
		t1	t2	t1	t2
Auditief	a	.91 (.06)	.88 (.10)	2596 (1171)	1063 (219)
	c	.90 (.06)	.87 (.07)	3108 (1101)	2487 (1122)
Visueel	v	.95 (.03)	.88 (.09)	2543 (800)	1173 (469)
	c	.93 (.10)	.94 (.04)	2868 (996)	2750 (707)
AV	av	.92 (.11)	.74 (.23)	2311 (599)	793 (374)
	c	.97 (.03)	.96 (.04)	2883 (693)	2735 (781)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training en c = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

Transfertaken

Op iedere taak is een 4 (training) × 2 (tijdstip) ANOVA uitgevoerd. Daarnaast zijn geplande contrasten berekend tussen één van de drie experimentele trainingen vs. de controletraining.

1. *EMT.* Er is een significant hoofdeffect van tijdstip en een significante training × tijdstip interactie ($F(3,74) = 3.14$, $p < .05$). Geplande contrasten tonen aan dat als gevolg van de training de controlegroep significant lager scoort dan de auditieve groep ($F(1,74) = 9.41$, $p < .01$) en de auditief-visuele groep ($F(1,74) = 3.81$, $p < .05$) en marginaal significant lager dan de visuele groep ($F(1,74) = 2.66$, $p = .10$). Zie Tabel 3.

2a. 200 ms. Flitstaak Met Woorden. Er is een significant hoofdeffect van tijdstip ($F(1,74) = 36.0$, $p < .001$). Alle groepen gaan in even grote mate significant vooruit. Over alle groepen heen gaat de vooruitgang van 35.5% goed gelezen woorden op de voortoets (*SD* = 23.5) naar 46.9% op de natoets (*SD* = 28.8).

2b. 200 ms. Flitstaak Met Pseudowoorden. Er is een significant hoofdeffect van tijdstip ($F(1,74) = 7.40$, $p < .01$). Alle groepen gaan in even grote mate significant vooruit. Over alle groepen heen gaat de vooruitgang van 59.4% goed gelezen pseudowoorden op de voortoets (*SD* = 21.5) naar 63.9% op de natoets (*SD* = 21.7).

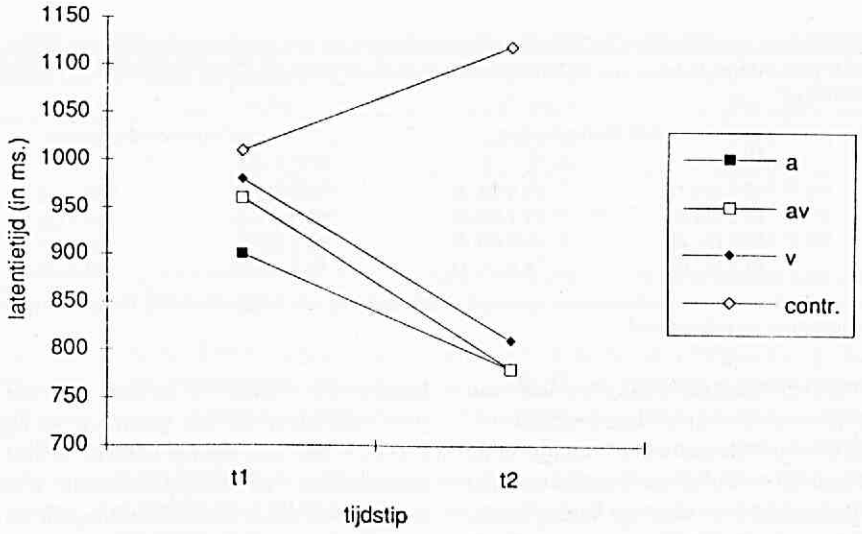
3a. Woorden Benoemen. Op het percentage goede antwoorden is er een hoofdeffect van tijdstip ($F(1,74) = 9.17$, $p < .01$). Alle groepen gaan na de training nauwkeuriger lezen: van 86.1% op de voortoets (*SD* = 11.8) naar 90% op de natoets (*SD* = 11.8). Ook op de latentietijden is er een significant hoofdeffect van tijdstip ($F(1,74) = 25.9$, $p < .001$). Alle groepen gaan op de natoets zo'n 400 ms. sneller lezen dan op de voortoets. Over alle groepen heen gaat de vooruitgang van 2460 ms. (*SD* = 1380) naar 2050 ms. (*SD* =1115).

3b. Pseudowoorden Benoemen. Op het percentage goede antwoorden is er een hoofdeffect van tijdstip ($F(1,74) = 3.85$, $p < .05$). Alle groepen gaan nauwkeuriger lezen: van 80.3% op de voortoets (*SD* = 15.7) naar 82.9% op de natoets (*SD* = 14.6). Op de latentietijden is er eveneens een hoofdeffect van tijdstip ($F(1,74) = 8.25$, $p < .001$). Over alle groepen heen wordt er op de natoets tussen de 100 en 300 ms. sneller gelezen dan op de voortoets. De gemiddelde vooruitgang gaat van 2410 ms. (*SD* = 1250) naar 2160 ms. (*SD* = 1290).

4. Letters Benoemen. Op het aantal gemaakte fouten zijn geen significante effecten gevonden. Over alle groepen en tijdstippen heen wordt er gemiddeld 1.5 fout (*SD* = 1.3) gemaakt. Op de latentietijden zijn alle hoofdeffecten significant en is er een significante training × tijdstip interactie ($F(3,74) = 5.12$, $p < .01$). Geplande contrasten tonen aan dat de controlegroep een significant hogere latentietijd heeft en dus langzamer presteert dan de auditieve (F=7.75), de visuele (F=11.5) en de auditief-visuele groep (F=12.1). Zie Figuur 1.

Effecten van getrainde subwoordeenheden

Er is een 4 (training) × 2 (tijdstip) × 2 (cluster: getrainde vs. ongetrainde) ANOVA uitgevoerd op de ruwe scores van 1) de detectietaken (*A'* scores en latentietijden) en 2) de gecomputeriseerde leestaken (percentage goede scores gesommeerd over flitstaken en benoemtaken, latentietijden alleen over benoemtaken). Er is sprake van een effect van getrainde subwoordeenheden omdat de geplande contrasten een significante training × tijdstip × cluster interac-



Figuur 1. Resultaten op de taak Letters Benoemen

Tabel 3
Scores op de EMT

Training	Tijdstip t1	Tijdstip t2
a	28.3 (10.5)	34.3 (11.9)
av	26.9 (9.10)	31.4 (11.0)
v	26.8 (10.0)	30.8 (9.84)
c	27.7 (7.34)	29.6 (7.87)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training en c = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

tie laten zien ten voordele van de getrainde clusters bij de experimentele training.

1. Detectietaken. Er is geen effect van getrainde subwoordeenheden. Net als bij de controlegroep worden bij de experimentele trainingsgroepen de getrainde clusters even snel en nauwkeurig herkend als de ongetrainde clusters.

2. Leestaken. Er is geen effect van getrainde clusters op de latentietijden. Op de percentage goede scores is er een significante training × tijdstip × cluster interactie bij geplande contrasten tussen de auditief-visuele en de controletraining ($F(1,307) = 5.39, p < .05$). De interactie is marginaal significant bij geplande contrasten tussen de visuele en de controletraining ($F(1,307) = 3.32, p = .07$). De interacties laten zien dat na de visuele en de auditief-visuele training, de woorden die een ongetraind cluster bevatten wel beter worden gelezen dan vóór de training maar dat de vooruitgang bij de woorden die een getraind cluster bevatten groter is. Zie Tabel 4.

Retentiemeting

Op iedere transfertaak zijn geplande contrasten berekend tussen de natoets (t2) en de retentietoets (t3) in een 4 (training) × 3 (tijdstip: t1 vs. t2 vs. t3) ANOVA. Het effect van getrainde clusters is berekend in een 4 (training) × 2 (cluster) ANOVA op de scores van de retentietoets.

Alleen op de EMT is er een significant verschil tussen de natoets en de retentietoets, ten voordele van de retentietoets ($F(1,142) = 20.0$, $p < .001$). Er zijn echter geen verschillen meer tussen de experimentele en de controletraining. De vooruitgang over alle trainingen heen gaat van 31.7 wpm. ($SD = 10.1$) op de natoets naar 34.7 wpm. ($SD = 11.2$) op de retentietoets. Op de overige leestaken zijn er geen vooruitgangen meer. De prestaties blijven op hetzelfde niveau als die op de eindtoets. Ook het effect van getrainde clusters bij de auditief-visuele en de visuele training verdwijnt in de retentiemeting.

1.4 Discussie

In antwoord op de vragen die in de inleiding van Experiment 1 gesteld worden vinden we de volgende resultaten. De eerste bevinding is dat de experimentele training generaliseert naar het sneller herkennen van subwoordeenheden in zijn algemeenheid. Dus zowel getrainde als ongetrainde clusters worden sneller in woorden herkend. Waarschijnlijk hebben de dyslectische kinderen als gevolg van de visuele en de auditief-visuele training een strategie aangeleerd waarbij meer risico's genomen worden

Tabel 4
Gemiddelde percentage goede scores bij woorden met getrainde en ongetrainde clusters op de Flitstaken en Benoemtaken

Training	Getrainde clusters		Ongetrainde clusters	
	t1	t2	t1	t2
a	65.8 (25.7)	70.6 (25.7)	64.2 (29.3)	69.0 (28.0)
av	66.0 (26.8)	75.3 (23.8)	61.1 (29.4)	66.3 (28.3)
v	63.8 (27.7)	70.5 (27.4)	63.3 (27.5)	67.6 (27.9)
c	71.8 (27.0)	74.5 (25.1)	65.5 (25.7)	72.2 (25.2)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training en c = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

(verlaging beslissingscriterium) ten bate van het verhogen van de verwerkingssnelheid.

Specifiek voor de subwoordtraining is de transfer naar de snelheid van letterherkenning en de transfer naar één vorm van hardop lezen, het zogenaamde hardop lezen van continue woorden, d.i. woorden die in een lijst gepresenteerd worden (EMT). Op beide vaardigheden is er namelijk een surplus van elke variant van de experimentele training ten opzichte van de controletraining. Ook is er transfer naar het hardop lezen van woorden die discreet, d.i. één voor één, gepresenteerd worden op het scherm. Die vooruitgang is echter niet specifiek voor de experimentele training omdat er geen verschillen zijn met de controletraining. De vooruitgangen op alle leestaken zijn toe te schrijven aan effecten van de experimentele training omdat het stopzetten van de training niet meer tot verdere vooruitgangen leidt. Alleen de prestaties op de EMT gaan nog vooruit op de retentietoets maar er zijn geen verschillen meer met de controletraining.

De derde bevinding is dat het oefenen van de letterklankstructuur leidt tot het toepassen van kennis van subwoordeenheden bij het lezen van woorden. Na de auditief-visuele training worden namelijk woorden met een getraind cluster beter gelezen dan woorden met een ongetraind cluster. Er is een tendens dat hetzelfde geldt voor de visuele training. Daarentegen is er geen evidentie dat dyslectische kinderen de fonologische kennisbron die ze aanleren in de auditieve training gebruiken bij het lezen van woorden.

Een aantal onderzoekers (e.g. Stanovich, 1981; Wolf, 1986) beweert dat een discrete leestaak, zoals de flitstaak en de benoemtaak, veel meer een beroep doet op de basale processen van woordherkenning terwijl een continue leestaak, zoals de EMT, ook nog allerlei hogere-orde strategieën toelaat, zoals het partieel verwerken van een woord en het tegelijkertijd verwerken van het volgend woord. Het surpluseffect van de experimentele training, ten opzichte van de controletraining, op de continue leestaak en niet op de discrete leestaken zou derhalve zo geïnterpreteerd kunnen worden dat de training eerder leidt tot een verbetering van de hogere-orde strategieën en niet zozeer tot een verbetering van de lagere orde processen, zoals het decoderen, de fonologische en de orthografische verwerking. Echter, er zijn bevindingen die juist het tegengestelde doen vermoeden. Het surpluseffect doet zich namelijk wel voor op de genoemde lagere orde processen op subwoordniveau (detectietaken) en de training leidt bovendien tot het toepassen van letter-klankeenheden bij het lezen.

Drie alternatieve verklaringen kunnen gegeven worden voor het uitblijven van differentiële effecten tussen de experimentele en de controletraining op de discrete leestaken. De trainingsduur is te kort, de moeilijkheidsgraad van de flitstaak is te hoog en, gegeven de complexiteit van het leesproces, zouden de drie varianten van de subwoordtraining wel eens te specifiek gericht kunnen zijn op de verbetering van slechts één aspect van woordkennis. In de controletraining worden immers veel meer alle basale aspecten van woordkennis getraind. Bovengenoemde verklaringen zullen in het volgende experiment getest worden.

2 Experiment 2

2.1 De vraagstelling

In het vorige experiment is aangetoond dat een training in het herkennen van subwoordeenheden ertoe leidt dat dyslectische kinderen sneller de beschikking hebben over subwoordeenhe-

Tabel 5
Leerlinggegevens: gemiddelden per experimentele en controle-conditie

Training	*N*	Leeftijd (mnd.)	EMT (wpm.)	PPVT (IQ)
a	9	122.4 (9.58)	22.9 (6.31)	103.6 (13.4)
av	11	120.9 (8.96)	22.7 (4.84)	110.0 (10.0)
v	10	121.6 (7.99)	27.1 (7.25)	106.0 (8.64)
comb.	10	123.4 (13.4)	28.3 (8.39)	107.1 (12.8)
contr.	10	123.1 (8.58)	26.0 (6.27)	100.5 (8.26)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training, comb. = combinatietraining en contr. = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

den in zijn algemeenheid. Dat wil zeggen dat niet alleen getrainde maar ook ongetrainde subwoordeenheden sneller herkend worden in woorden. Daarbij moet opgemerkt worden dat de generalisatie beperkt blijft tot de getrainde kennisstructuur. In Experiment 2 zal nagegaan worden of een subwoordtraining in een bepaalde kennisstructuur ook tot transfer leidt naar andere kennisstructuren. Leidt bijvoorbeeld een training in fonologische subwoordeenheden tot een verbetering in de herkenning van orthografische subwoordeenheden? Anders gezegd, leidt versterking van één kennisbron tot versterking van andere kennisbronnen?

In het eerste experiment is gebleken dat de transfer naar het hardop lezen van discrete woorden niet specifiek was voor de subwoordtraining omdat de controletraining tot even grote vooruitgang leidde. In het tweede experiment zal opnieuw het effect van de subwoordtraining onderzocht worden maar nu onder andere condities. De trainingsduur wordt verdubbeld naar 320 minuten. Om te controleren voor de moeilijkheidsgraad van de flitstaak worden nu twee aanbiedingstijden ingesteld, een relatief korte (200 ms.) en een relatief lange (500 ms.). Bovendien wordt er een vierde variant van de experimentele training toegevoegd, de combinatietraining. In de combinatietraining worden zowel orthografische als fonologische subwoordeenheden als de combinatie daartussen in één training geoefend.

Tot slot wordt in Experiment 2 nader onderzocht of de subwoordtraining effect heeft op de basale processen of op de hogere orde processen van lezen. Deze vraag wordt op twee manieren onderzocht. Op de eerste plaats wordt er een onderscheid gemaakt in transfer op continue en discrete leestaken. Het lezen van continue woorden zou meer een beroep doen op de hogere processen, het lezen van discrete woorden meer op de basale processen (e.g. Stanovich, 1981; Wolf, 1986). In vergelijking tot het vorig experiment wordt hiertoe het assortiment van leestaken vergroot. Op de tweede plaats wordt er op het niveau van de discrete leestaken een onderscheid gemaakt in transfer naar het lezen van woorden en het lezen van pseudowoorden. Het lezen van woorden doet meer een beroep op de hogere processen omdat gebruik gemaakt wordt van de bekendheid en de betekenis. Het lezen van pseudowoorden, die geen betekenis en voordeel van bekendheid hebben, doet vooral een beroep op het basale proces van decoderen.

Samengevat moet Experiment 2 antwoord geven op de volgende vragen.

1. Leidt een subwoordtraining in één basaal proces van woordlezen tot verbetering van andere basale processen?
2. Is de combinatietraining, waarin alle basale aspecten van woordkennis getraind worden, effectiever dan de controletraining?
3. Zijn er verschillen in transfer naar hogere en lagere orde leesprocessen?

a. Is er verschil in transfer op continue en discrete leestaken?

b. Is er verschil in transfer naar het lezen van woorden en pseudowoorden?

2.2 Methode

Proefpersonen

Vijftig zeer zwakke lezers uit het LOM-onderwijs doen mee aan het onderzoek, verdeeld over vijf trainingen. Net als in Experiment 1 zijn de groepen gematcht op leeftijd en technische leesvaardigheid (EMT) en hebben ze een gelijke verbale intelligentie (PPVT IQ). Zie Tabel 5.

Algemene procedure

Het design is een pretest-training-posttest design met een retentiemeting. De volgorde van afnames is: voortoets (t1) - eerste trainingsblok - tussentoets (t2) - tweede trainingsblok - eindtoets (t3) - retentietoets (t4). Elk trainingsblok duurt 160 minuten. De totale trainingstijd is dus 320 minuten, opgedeeld in 32 sessies van 10 minuten. Na iedere 16 sessies is er een toets afgenomen ter evaluatie van de leesprestatie. De tijd tussen de voor- en de tussentoets is even lang als die tussen de tussen- en de eindtoets, en bedraagt twee maanden. De tijd tussen de eind- en de retentietoets is drie maanden. In alle toetsen worden alle leestaken gebruikt, behalve de Pseudo-EMT (zie toetsen). Die wordt alleen in de eind- en de retentietoets afgelegd. De Detectietaken worden alleen bij de eindtoets gebruikt.

Trainingstaken

Zie Experiment 1 voor een beschrijving van de trainingstaken, inclusief de controletraining. Nieuw in dit experiment is de combinatietraining. De combinatietraining bestaat uit 10 sessies auditieve training, vervolgens 11 sessies visuele training en tot slot 11 sessies auditief-visuele training. De combinatietraining duurt dus even lang als de overige trainingen en beslaat 32 sessies. In 32 sessies komt ieder cluster 16 keer aan bod. Zie verder Experiment 1 voor de procedure van de training.

Toetsen

A. Detectietaken

Zie Experiment 1 voor een beschrijving van de drie varianten van de taak, te weten de *Auditieve-*, *Visuele-* en *Auditief-Visuele Detectietaak*. Deze taken worden gebruikt om na te gaan of de subwoordtraining leidt tot transfer naar een ongetraind aspect van woordkennis. Daarom worden, na de totale trainingstijd, alle trainingsgroepen gevraagd alle varianten van de taak uit te voeren.

B. Continue leestaken

1. *Eén-Minuut Test* (EMT). Zie Experiment 1.

2. *Pseudo-EMT*. Analoog aan de EMT maar nu met pseudowoorden. Deze taak wordt alleen in de eind- en de retentietoets gebruikt.

3. *Differentiële Zinnen leesTest* (DZT, versie A; Dommerholt, 1969). Kinderen moeten in drie minuten tijd zo goed en zo veel mogelijk zinnen lezen. De test bestaat uit drie sets van elk 10 zinnen. Per set neemt de moeilijkheidsgraad van de zinnen toe. De score is het aantal goed gelezen woorden per minuut.

C. Discrete leestaken

1. *Flitstaken*. Er zijn 4 flitstaken: a) 200 ms. Flitstaak Met Woorden: zie Experiment 1, b) 500 ms. Flitstaak Met Woorden: analoog aan a) maar nu met een aanbiedingstijd van 500 ms, c) 200 ms. Flitstaak Met Pseudowoorden: zie Experiment 1, en d) 500 ms. Flitstaak Met Pseudowoorden: analoog aan c) maar nu met een aanbiedingstijd van 500 ms.

2. *Benoemtaken*. Zie Experiment 1: a) Woorden Benoemen, b) Pseudowoorden Benoemen.

2.3 Resultaten

Detectietaken

Op iedere detectietaak zijn geplande contrasten berekend tussen een variant van de experimentele training vs. de controletraining op *A'* scores en latentietijden. In Tabel 6 zijn de scores weergegeven.

1. Auditieve Detectietaak. Zonder af te doen aan de nauwkeurigheid, presteert de auditieve en de combinatiegroep significant sneller dan de controlegroep ($F(1,44)$ = 15.5 resp. 13.3, $p < .001$) en de auditief-visuele groep marginaal significant sneller ($F(1,44) = 3.47$, $p = .07$). De visuele groep heeft daarentegen eenzelfde latentietijd als de controlegroep en maakt zelfs meer fouten ($F(1,45) = 4.82$, $p < .05$).

2. *Visuele Detectietaak.* Niet alleen de visuele en de combinatiegroep presteert significant sneller dan de controlegroep ($F(1,45)$ = 45.8 resp. 50.6, $p < .001$), ook de auditieve en de auditief-visuele groep presteren sneller (F = 18.5 resp. $F = 52.1$). Wel maken de laatste twee groepen, die niet rechtstreeks getraind zijn in de orthografische vaardigheid die deze taak vergt, meer fouten dan de controlegroep (F = 8.77 resp. $F = 9.42$).

3. Auditief-Visuele Detectietaak. Niet alleen de auditief-visuele en de combinatiegroep presteert sneller dan de controlegroep ($F(1,45)$ =52.6 resp. 50.6, $p < .001$), ook de auditieve en

Tabel 6
Scores op de Detectietaken in de eindtoets

Taak	Training				
	a	av	v	comb.	contr.
Auditief					
A'	.85 (.10)	.85 (.09)	.78 (.16)	.88 (.08)	.86 (.07)
latentietijd (in ms.)	1136 (443)	1823 (1049)	2150 (757)	1262 (459)	2383 (716)
Visueel					
A'	.79 (.19)	.80 (.12)	.86 (.14)	.89 (.08)	.94 (.07)
latentietijd (in ms.)	1741 (962)	1063 (585)	1136 (432)	1047 (343)	2878 (504)
AV					
A'	.84 (.17)	.90 (.09)	.89 (.10)	.90 (.14)	.93 (.05)
latentietijd (in ms.)	1504 (698)	923 (512)	1360 (644)	916 (250)	2619 (609)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training, comb. = combinatietraining en contr. = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

de visuele groep, die niet rechtstreeks getraind zijn in de decodeervaardigheid die deze taak vergt, presteren sneller (F = 20.6 resp. F = 27.6). Daarbij maakt de auditieve trainingsgroep marginaal significant meer fouten dan de controlegroep ($F = 3.73, p = .06$).

Samengevat, ongeacht welke kennisstructuur getraind wordt, leidt de experimentele training tot een snellere herkenning van orthografische en letter-klankclusters. Bij de groepen die niet rechtstreeks getraind zijn in de herkenning van visuele clusters gaat de snelheid ten koste van de nauwkeurigheid. Een verbeterde herkenning van fonologische subwoordeenheden blijft daarentegen beperkt tot die groepen die rechtstreeks getraind zijn met fonologische clusters.

Leestaken

Om effecten te berekenen na de totale trainingstijd van 320 minuten is een 5 (training) × 3 (tijdstip: voor- vs. tussen- vs. eindtoets) ANOVA uitgevoerd op de ruwe scores. Dit geldt voor iedere leestaak behalve de Pseudo-EMT; deze taak is namelijk alleen in de eindtoets gebruikt. Daarnaast zijn twee aparte 5 (training) × 2 (tijdstip) ANOVA's uitgevoerd om de effecten van het eerste trainingsblok (t1 vs. t2) te onderscheiden van de effecten van het tweede trainingsblok (t2 vs. t3). Bij iedere ANOVA zijn geplande contrasten berekend tussen één van de vier experimentele condities vs. de controletraining. Op één taak (500 ms. Flitstaak Met Woorden) waren er op de voortoets al significante verschillen tussen groepen. Er is hiervoor statistisch gecorrigeerd door in de ANOVA verschilscores tussen twee tijdstippen als afhankelijke variabelen te gebruiken.

Continue leestaken

1. EMT. Er is een significant hoofdeffect van tijdstip ($F(2,90) = 58.3, p < .001$), zowel in het eerste trainingsblok ($F(1,45) = 47.0, p < .001$) als in het tweede ($F(1,45) = 13.7, p < .001$). Een 5 (training) × 2 (tijdstip: t1 vs. t2) ANOVA laat een significante training × tijdstip interactie zien bij geplande contrasten tussen de combinatie en de controlegroep ($F(1,45) = 4.24, p < .05$). De interactie toont aan dat in de eerste trainingsperiode de kinderen uit de combinatiegroep sneller vooruit gaan dan de kinderen uit de controlegroep. Deze voorsprong verdwijnt echter na een verdubbeling van de trainingstijd want na de totale trainingstijd is de interactie niet meer significant; de controlegroep maakt een relatief sterke vooruitgang in het tweede trainingsblok (t2 vs. t3). Zie Tabel 7.

Tabel 7
Scores op de EMT

Training	Tijdstip		
	t1	t2	t3
a	22.9 (6.31)	26.0 (5.79)	27.4 (6.65)
av	22.7 (4.84)	27.0 (6.31)	29.9 (7.29)
v	27.1 (7.25)	31.3 (8.19)	32.4 (9.38)
comb.	28.3 (8.40)	33.9 (9.12)	34.9 (8.43)
contr.	26.0 (6.27)	28.0 (7.32)	30.9 (7.13)

Noot: a = auditieve training, av = auditief-visuele training, v = visuele training, comb. = combinatietraining en contr. = controletraining. De getallen tussen haakjes zijn standaarddeviaties.

2. Pseudo-EMT. Geplande contrasten op de scores van de eindtoets laten een significant

verschil zien tussen de combinatie en de controlegroep, ten voordele van de eerste ($F(1,45) = 4.45, p < .05$). De combinatiegroep leest 20.5 pseudowoorden per minuut correct ($SD = 7.11$) terwijl de controlegroep slechts een score van 15.4 heeft ($SD = 5.19$). De overige experimentele groepen scoren even hoog als de controlegroep. Het gemiddelde van de auditieve groep is 14.3 (SD =4.00), van de auditief-visuele groep 16 ($SD = 3.69$) en van de visuele groep 16.8 ($SD = 6.27$).

3. DZT. Er is een significant hoofdeffect van tijdstip ($F(2,90) = 35.2, p < .001$), zowel in het eerste trainingsblok ($F(1,45) = 15.1, p < .001$) als in het tweede ($F(1,45) = 25.0, p < .001$). De vooruitgang op het lezen van woorden in zinnen is bij alle groepen gelijk. Over alle groepen heen is de gemiddelde score 45.8 wpm. op t1 ($SD = 11.1$), 49.5 wpm. op t2 ($SD = 10.2$) en 53.8 wpm. op t3 ($SD = 11.9$).

Discrete leestaken

1a. 200 ms. Flitstaak Met Woorden. Er is een hoofdeffect van tijdstip ($F(2,90) = 57.3, p < .001$), zowel in het eerste trainingsblok ($F(1,45) = 51.3, p < .001$) als in het tweede ($F(1,45) = 17.3, p < .001$). De vooruitgang is bij alle groepen even groot. Het gemiddeld percentage goed gelezen woorden is 31.5 op t1 ($SD = 20.9$), 48.2 op t2 ($SD = 26.0$) en 59.0 op t3 ($SD = 27.1$).

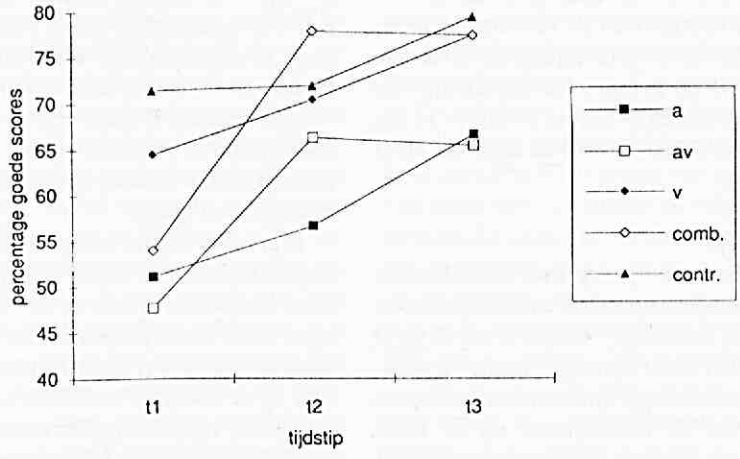
1b. 500 ms. Flitstaak Met Woorden. Zie Figuur 2a. Uit geplande contrasten blijkt dat de controlegroep op de voortoets significant beter scoort dan de auditief-visuele groep ($F(1,45) = 5.19, p < .05$) en marginaal significant beter dan de auditieve groep $F(1,45) = 3.95, p = .07$). Ter correctie worden verschilscores gebruikt. In de eerste trainingsperiode (t2 - t1) presteren de auditief-visuele en de combinatiegroep beter dan de controlegroep ($F(1,45) = 6.70, p < .01$ resp. $F(1,45) = 11.2, p < .01$). In het tweede trainingsblok (t3 - t2) zijn er geen significante verschillen tussen groepen. Na de totale trainingstijd (t3 - t1) presteert alleen de combinatiegroep beter dan de controlegroep ($F(1,45) = 4.95, p < .05$). De vooruitgang bij de kinderen uit de combinatiegroep is dus voornamelijk te danken aan het effect van de eerste trainingsperiode. In deze periode hebben zij alleen nog maar de auditieve en de visuele training gehad en nog niet de auditief-visuele training.

1c. 200 ms. Flitstaak Met Pseudowoorden. Zie Figuur 2b. Er is een significant hoofdeffect van tijdstip ($F(2,90) = 14.4, p < .001$) die ook optreedt in het eerste trainingsblok ($F(1,45) = 7.27, p < .01$) en in het tweede ($F(1,45) = 7.79, p < .01$). Geplande contrasten na de totale trainingstijd tonen aan dat alleen de combinatiegroep ten opzichte van de controlegroep vooruitgaat ($F(2,90) = 5.63, p < .01$). Deze vooruitgang is het gevolg van het tweede trainingsblok (t3 vs. t2) want alleen in deze periode, en niet in de eerste, is de vooruitgang significant ($F(1,45) = 6.72, p < .05$). In de tweede trainingsperiode heeft de combinatiegroep naast de auditieve en visuele training ook de auditief-visuele training gedaan. In de eerste trainingsperiode (t2 vs. t1) zijn de vooruitgangen bij alle groepen even groot.

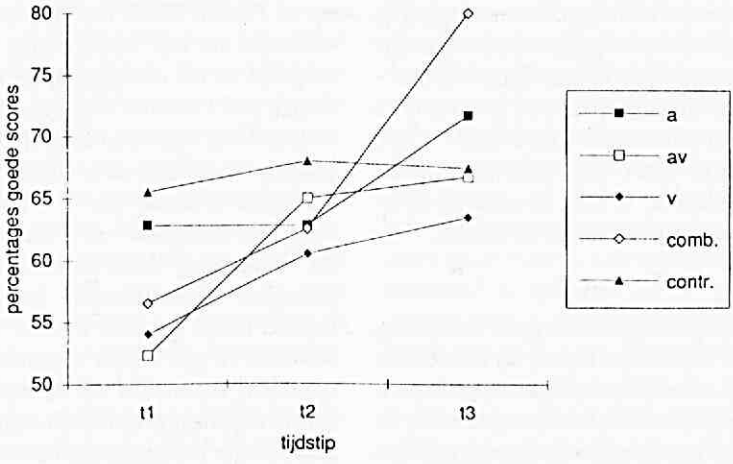
1d. 500 ms. Flitstaak Met Pseudowoorden. Op deze taak zijn geen significante effecten gevonden. Het gemiddeld percentage goede scores is 72.2 op t1 ($SD = 19.0$), 74.2 op t2 ($SD = 15.5$) en 76.3 op t3 ($SD = 19.4$).

2a. Woorden Benoemen. Op het percentage goede antwoorden zijn geen significante effecten gevonden. Over alle tijdstippen en groepen is het gemiddeld percentage goed gelezen woorden 90.5% ($SD = 10.7$). Op de latentietijden is er een hoofdeffect van tijdstip ($F(2,90) = 28.4, p < .001$) die zowel optreedt in de eerste trainingsperiode ($F(1,45) = 25.6, p < .001$) als in de tweede ($F(1,45) = 9.35, p < .05$). Geplande contrasten na de totale trainingstijd tonen aan dat alleen de combinatiegroep ten opzichte van de controlegroep vooruitgaat ($F(2,90) = 3.27, p < .05$). Zie Figuur 2c. Deze vooruitgang is voornamelijk het effect van het eerste trainingsblok (t2 vs. t1) want alleen in deze periode, en niet in de tweede, is de vooruitgang marginaal significant ($F(1,45) = 3.33, p = .07$). In de tweede trainingsperiode (t3 vs. t2) zijn de vooruitgangen bij alle groepen even groot.

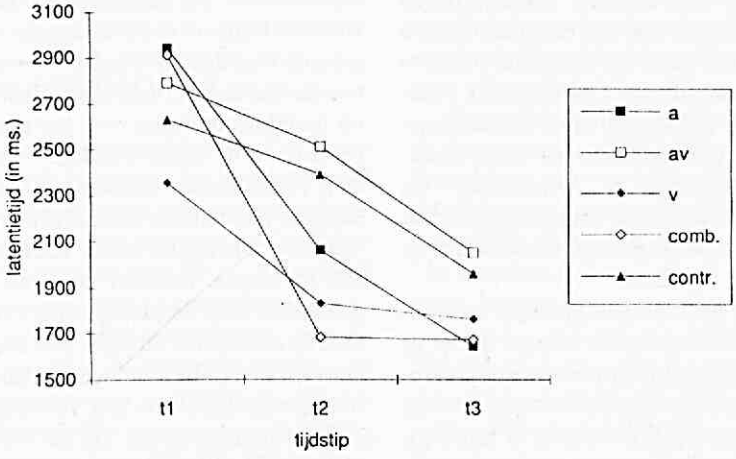
2b. Pseudowoorden Benoemen. Op het percentage goede antwoorden zijn geen significante effecten gevonden. Over alle tijdstippen en groepen is het gemiddeld percentage goed gelezen pseudowoorden 84.2% (SD = 13.6). Op de latentietijden is er een hoofdeffect van tijdstip ($F(2,90) = 11.5, p < .001$) die ook te zien is in de eerste trainingsperiode ($F(1,45) = 10.5, p < .01$) maar niet in de tweede ($F(1,45) =$



Figuur 2a. Resultaten op de 500 ms. Flitstaak Met Woorden



Figuur 2b. Resultaten op de 200 ms. Flitstaak Met Pseudowoorden



Figuur 2c. Resultaten op de taak Woorden Benoemen

1.85. $p > .15$). Alle groepen gaan dus in even grote mate vooruit maar de vooruitgang is beperkt tot de eerste trainingperiode. Over alle trainingen is de gemiddelde latentietijd van pseudowoorden 2610 ms. op t1 (SD = 1110), 2250 ms. op t2 (SD = 1260) en 2120 ms. op t3 (SD = 1100).

Retentietoets

Op iedere leestaak zijn geplande contrasten berekend tussen de natoets (t3) en de retentietoets (t4) in een 5 (training) × 4 (tijd: t1 vs. t2 vs. t3 vs. t4) ANOVA. Er zijn significante verschillen tussen de natoets en de retentietoets, ten voordele van de retentietoets, op de EMT ($F(1,132) = 6.15$, $p < .001$) en de DZT ($F(1,132) = 7.85$, $p < .001$). Er zijn geen verschillen tussen de trainingen. De vooruitgang op de EMT gaat van 31.0 wpm. in de eindtoets (SD = 7.90) naar 32.3 wpm. in de retentietoets (SD = 8.48). De vooruitgang op de DZT gaat van 53.8 wpm. (SD = 11.9) naar 56.1 wpm. (SD = 11.1). Op de overige taken zijn geen significante vooruitgangen meer te zien; de score op de retentietoets is gelijk aan die op de natoets.

2.4 Discussie

In antwoord op de eerste vraag, die betrekking heeft op de interacties tussen de drie basale bronnen van woordkennis, blijkt uit het experiment dat, wat betreft subwoord-eenheden, de fonologische kennisbron bij dyslectische kinderen het minst te beïnvloeden is via versterking van de overige bronnen. Alleen groepen die rechtstreeks getraind zijn met fonologische clusters kunnen namelijk dergelijke clusters sneller herkennen dan de controlegroep. Daarentegen is een verbetering in de herkenningssnelheid van orthografische en letter-klankclusters gemakkelijk te beïnvloeden via versterking van de andere bronnen, zij het dat dit enigszins ten koste gaat van de nauwkeurigheid.

De tweede bevinding is dat er op verschillende hardop-leestaken een surplus is van de combinatietraining ten opzichte van de controletraining. Het blijkt dat voor een transfer naar bestaande woorden een training in fonologische en orthografische subwoordeenheden voldoende is (auditieve resp. visuele training), dus zonder extra training in letter-klankkoppelingen (zie Tabel 7, en Figuur 2a en c). Daarentegen lijkt een training in letter-klankkoppelingen wel noodzakelijk te zijn voor een transfer naar pseudowoorden. Het surplus op de Pseudo-EMT en de 200 ms. Flitstaak Met Pseudowoorden (Figuur 2b) treedt namelijk pas op nadat de combinatiegroep naast de auditieve en visuele training ook de auditief-visuele training heeft gehad.

In antwoord op de vraag of de subwoordtraining de lagere of de hogere orde processen van lezen stimuleert vinden we de volgende resultaten. In de eerste plaats doet het surplus op de discrete leestaken (Figuur 2a, b en c) en niet zozeer op de continue leestaken, vermoeden dat de training vooral de lagere, basale leesprocessen beïnvloedt. Op de continue leestaken is het surplus na de totale trainingstijd slechts te zien op de Pseudo-EMT. Het surplus op de EMT verdwijnt na een verdubbeling van de trainingstijd en de vooruitgang op het lezen van zinnen (DZT) is niet verschillend van de controletraining. Het is niet duidelijk of de vooruitgang op deze laatste twee taken te wijten is aan specifieke effecten van de subwoordtraining omdat het stopzetten van de training nog steeds tot vooruitgang leidt, in tegenstelling tot de prestaties op de discrete leestaken. In de tweede plaats doet de transfer naar discrete woorden en niet zozeer naar discrete pseudowoorden, vermoeden dat de training niet zozeer het basale proces van decoderen stimuleert maar eerder processen die interacteren met de hogere orde verwerking van betekenisaspecten van woorden. De gemakkelijke transfer naar woorden blijkt uit de bevinding dat de vooruitgang op woordtaken zowel in de eerste als in de tweede trainingsperiode plaatsvindt terwijl die op pseudowoordtaken over het algemeen beperkt is tot de eerste trainingsperiode. Op de taak Pseudowoord Benoemen blijkt bijvoorbeeld na 160 minuten training al een asymptoot bereikt te zijn die tot aan de retentietoets over alle trainingen, inclusief de controletraining, gemiddeld rond de 2180 ms. ligt (SD = 1070). Ook op de 200 ms. Flitstaak Met Pseudowoorden zien we bij alle trainingen, behalve de combinatietraining, alleen een vooruitgang in de eerste trainingsperiode. Op de 500 ms. Flitstaak Met Pseudowoorden is er zelfs helemaal geen vooruitgang.

Een aantal vragen blijft na dit onderzoek onbeantwoord en dient nader onderzocht te wor-

den. Waarom verdwijnt bijvoorbeeld het surplus van de combinatietraining op de EMT na een langere trainingstijd? Zegt dit meer iets over de controletraining of over de experimentele training? De controletraining kan immers omschreven worden als een minder specifieke, maar wel veelzijdiger methode die in elk geval effect heeft bij zwakke lezers die minder achterstand vertonen dan de leerlingen in het onderhavige experiment (Smeets & Van der Leij, 1993). Wellicht is zo'n methode op den duur ook effectief bij dyslectische leerlingen. Een publikatie hieromtrent is in voorbereiding. Een vergelijking met een controlegroep die helemaal geen leestraining krijgt kan bovendien in een volgend onderzoek mogelijkheden bieden om te controleren op de factor 'intensieve training'. Een volgende vraag is waarom de aanbiedingstijd op de Flitstaken aanleiding geeft tot verschillende resultaten. Waarschijnlijk is het lezen van tweelettergrepige woorden die slechts 200 ms. te zien zijn te moeilijk, in vergelijking tot de 500 ms. conditie, om subtiele verschillen te vinden tussen de experimentele en de controletraining. Bij het lezen van pseudowoorden zou een plafondeffect in de 500 ms. conditie kunnen verklaren waarom er geen vooruitgangen te zien zijn in deze conditie en wel in de 200 ms. conditie. Een andere verklaring is dat dyslectische lezers bij het lezen van pseudowoorden gestimuleerd worden om efficiëntere strategieën toe te passen bij een korte aanbiedingstijd dan bij een lange, en dat de combinatiegroep daar het meest in geslaagd is.

3 Algemene discussie

In twee experimenten is nagegaan of een computergestuurde training in het snel herkennen van subwoordeenheden leidt tot een verbetering van de processen die ten grondslag liggen aan snel en efficiënt lezen. De resultaten tonen aan dat de training leidt tot een snellere verwerking van letters, subwoordeenheden, woorden en pseudowoorden. Wanneer alle aspecten van woordkennis getraind worden is er zelfs een surplus ten opzichte van een controletraining waarin eveneens intensief met woorden geoefend wordt op de computer. De subwoordtraining heeft vooral effect op het lezen van bestaande woorden die geïsoleerd worden aangeboden. Dit geeft aan dat de training vooral die basale processen van woordherkenning stimuleert die interacteren met de bekendheid met betekenis van de woorden.

De resultaten hebben implicaties voor de ontwikkeling van woorherkenningsmodellen. In de inleiding is het belang van de verwerking van letterclusters aangegeven volgens connectionistische principes. In hoeverre beantwoorden de prestaties van dyslectische lezers nu aan de assumpties van het connectionistisch model? In de eerste plaats steunen de resultaten het idee dat de vorming van letterclusters gebaseerd is op een snelle herkenning van afzonderlijke letters. Het sneller herkennen van letterclusters gaat namelijk gepaard met het sneller herkennen van individuele letters. In de tweede plaats steunen de resultaten het idee dat subwoordeenheden de basiselementen zijn van woordherkenning. Een training in het herkennen van subwoordeenheden leidt namelijk tot een snellere herkenning van woorden. Bovendien is er evidentie dat bij het lezen van woorden, kennis van de letter-klankstructuur van subwoordeenheden wordt toegepast. In de derde plaats beperken de resultaten het connectionistisch leerprincipe dat de versterking van één kennisbron vanzelf leidt tot een verbetering van de andere bronnen. Uit het onderzoek blijkt namelijk dat er bij dyslectische lezers weinig transfer is naar de fonologische kennisbron terwijl er wel een transfer is naar de bronnen van orthografische- en de letter-klank kennis, onafhankelijk van welke kennisstructuur getraind wordt. Het is echter nog maar de vraag of dit resultaat ook te generaliseren is naar een populatie van normale lezers.

De resultaten zijn in overeenstemming met de hypothese dat dyslectische leerlingen een fonologisch tekort vertonen (Snowling, 1987). De geringe transfer naar de herkenning van fonologische subwoordeenheden en het lezen van pseudowoorden geeft aan dat het fonologisch probleem zo hardnekkig is dat het zelfs met gerichte instructie moeilijk te verhelpen is. Het feit dat er een relatief gemakkelijke transfer is naar de herkenning van orthografische subwoordeenheden en het lezen van bestaande woorden doet vermoeden dat dyslectische lezers hun fonologische zwakte compenseren met kennis van de orthografie, betekenis en bekendheid van woorden. De notie van compen-

satie is de keerzijde van de tekorthypothese (zie o.a. Nicolson & Fawcett, 1990; Rack, Snowling & Olson, 1992). Juist deze combinatie tussen een fonologisch tekort en compensatiemogelijkheden via orthografie en betekenis zou in de behandeling van dyslexie centraal moeten staan.

Het onderzoek heeft ook een practische waarde. Aangezien de computergestuurde subwoordtraining, in de vorm van de combinatietraining, effectief is kan de training gebruikt worden in het leesonderwijs. Het algemeen voordeel van een computergestuurde training is dat het lezen individueel geoefend kan worden in de klas. Het specifieke voordeel van deze training is dat het kind leert om naast de accuratesse ook rekening te houden met de snelheid van werken. In bredere zin betekent het onderzoek dat het onderwijs in de basisvaardigheden van het lezen niet alleen hoeft plaats te vinden via de tot nu gangbare methode van het actief hardop verklanken van woorden. Via het articuleren van woorden wordt het kind doorgaans geconfronteerd met het besef dat het lezen van woorden een serieel proces is waarin letters en lettereenheden in de juiste volgorde verklankt moeten worden (Adams, 1990). Dit onderzoek impliceert dat de seriële verwerking in het leesproces ook geoefend kan worden door het kind orthografische en verklankte subwoordeenheden te laten identificeren in woorden. Doordat de eenheden telkens op verschillende posities in een woord voorkomen leert het kind dat woorden uit elementen bestaan die in een bepaalde volgorde staan. Daardoor wordt meer simultane verwerking van eenheden boven letterniveau toegevoegd aan de seriële verwerking. Door de druk op de verwerkingssnelheid leert het kind bovendien dat de aandacht efficiënt verdeeld moet worden over een woord.

Literatuur

Adams, M.J. (1990). *Beginning to read. Thinking and learning about print.* Cambridge, MA: MIT Press.

Baddeley, A.B. (1986). *Working memory.* Oxford: Clarendon Press.

Bosch, K. van den (1991). *Poor readers' decoding skills.* Effects of training, task and word characteristics. Academisch proefschrift Katholieke Universiteit Nijmegen.

Bouma, H., & Legein, C.P. (1980). Dyslexia: A specific recoding deficit? An analysis of response latencies for letters and words in dyslectics and in average readers. *Neuropsychologia, 18,* 285-298.

Bowers, P.G., & Swanson, L.B. (1991). Naming speed deficits in reading disability: multiple measures of single processes. *Journal of Experimental Child Psychology, 51,* 195-219.

Brus, B.Th., & Voeten, M.J.M. (1973.) *Eén-minuut-Test, vorm A en B.* Nijmegen: Berkhout.

Daal, V.H.P. van (1993). *Computer-based Reading and Spelling Practice for Young Dyslexics.* Academisch proefschrift Vrije Universiteit Amsterdam.

Daal, V.H.P. van, & Leij, A. van der, Bakker, N.C.M., Reitsma, P. (1987). Een computergestuurd orthodidactisch programma voor aanvankelijk lezen: COPAL. *Pedagogische Studiën, 64,* 364-376.

Dencla, M. & Rudel, R. (1976). Naming of object-drawings by dyslexic and other learning disabled children. *Brain and Language, 3,* 1-15.

Dommerholt, I. (1969.). *Differentiële zinnenleestest.* Groningen: Wolters-Noordhoff.

Dunn, D. & Lloyd, M. (1965). *Peabody Picture Vocabulary Test.* Washington: American Guidance Service Inc.

Ehri, L.C., & Wilce, L.S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75,* 3-18.

Frederiksen, J.R. (1982). A componential theory of reading skills and their interactions. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence. 1*: 125-180. Hillsdale, NJ: Lawrence Erlbaum Associates.

Frederiksen, J.R., Warren, B.M; Roseberry, A.S. (1985). A componential approach to training reading skills: Part 1. Perceptual units training. *Cognition and Instruction, 2*(2), 91-130.

Klapwijk, M.J.G., & Das- Smaal, E. (1990). *Leesvaardigheidstraining volgens een componentenbenadering.* Amsterdam: Vrije Universiteit.

LaBerge, D., Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Leij, A. van der (1990). De constructie van een computergestuurd programma voor aanvankelijk lezen. *Tijdschrift voor Orthopedagogiek, 5,* 245-257.

Lovett, M.W. (1987). A developmental approach to reading disability: Accuracy and speed criteria of normal and deficient reading skill. *Child Development, 58,* 234-260.

Manis, F. R. (1985). Acquisition of word identification skills in normal and disabled readers. *Journal of Educational Psychology, 77*, 78-90.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research? *Cognition, 35*, 159-182.

Pollack, I., Norman, D. A. (1964). A non-parametic analysis of recognition experiments. *Psychonomic Science, 1*, 125-126.

Rack, J. P., Snowling, M. J. & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: a review. *Reading Research Quarterly, 27*, 29-53.

Seidenberg, M. S., McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Smeets, H., & Leij, A. van der (1993). Differentiële effecten van computergestuurde instructie op het lezen van dyslectische en zwak lezende kinderen: een voorstudie. *Pedagogische Studiën, 70*, 56-72.

Snowling, M. (1987). *Dyslexia: A cognitive developmental perspective*. Oxford: Basil Blackwell Ltd.

Spring, C. & Davis, J. M. (1988). Relations of digit naming speed with three components of reading. *Applied Psycholinguistics, 9*, 315-334.

Stanovich, K. E. (1981). Relationships between word decoding speed, general name retrieval ability, and reading progress in first-grade children. *Journal of Educational Psychology, 73*, 809-815.

Wolf, M. (1986). Rapid alternative naming in the developmental dyslexias. *Brain and Language, 27*, 360-379.

Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes and reading: A longitudinal study in average and impaired readers. *Child Development, 57*, 988-1005.

Yap, R. (1993). *Automatic word processing defecits in dyslexia: qualitative differences and specific remediation.* Academisch proefschrift Vrije Universiteit Amsterdam.

Yap, R. L., & Leij, A. van der, (1993). Word processing in dyslexics: an automatic decoding deficit? *Reading and Writing: An Interdisciplinary Journal, 5*, 261-279.

Manuscript aanvaard 8-7-1993

Auteurs

R. Yap studeerde theoretische psychologie en neuropsychologie aan de Katholieke Universiteit Brabant. Van 1989 tot 1993 is zij werkzaam geweest als assistent in opleiding bij de sectie Speciale Pedagogiek van de Vrije Universiteit. Zij promoveerde in 1993 op een proefschrift waar de beschreven experimenten deel van uitmaakten.

A. van der Leij is als hoogleraar verbonden aan de sectie Speciale Pedagogiek van de Vrije Universiteit.

Correspondentie-adres: Vrije Universiteit, sectie Speciale Pedagogiek, Van der Boechorststraat 1, 1081 BT Amsterdam

Abstract

Computer based remediation of reading disability by increasing the recognition speed of subword units

R. Yap & A. van der Leij. Pedagogische Studiën, 1993, 70, 402-419.

In two experiments, the effect of a computer assisted trainingprogram at the sublexical level is investigated on several reading processes. Subjects are 10-year-old dyslexic readers lagging two years or more behind in reading. The training is aimed at improving recognition speed of orthographic or phonological subword units or a combination of both. Main results indicate that the training results in faster recognition of isolated letters, sublexical units, words, and pseudowords. When all three kinds of units are trained, the sublexical training is more effective than a control training in which words were also practiced on the computer. The training particularly improves reading of isolated words.

Training van lom-leerkrachten in directe instructie van begrijpend leesstrategieën*

S. E. Sliepen en P. Reitsma**

Samenvatting

Recent onderzoek toont aan dat in het reguliere basisonderwijs nauwelijks directe, expliciete instructie in leesstrategieën wordt gegeven. Over de situatie in het speciaal onderwijs is nog relatief weinig bekend. In dit onderzoek zijn daarom bij groepen in de lom-bovenbouw de instructievorm en -inhoud van een groot aantal lessen in begrijpend lezen op video-/audioband vastgelegd en gecodeerd met behulp van een observatie-categorieënsysteem en/of een waarderingschaal. Uit de resultaten blijkt dat ook in het lom-onderwijs doorgaans weinig directe instructie wordt gegeven in begrijpend lezen.

Vervolgens is getracht via een training van leerkrachten verandering te brengen in het lesgedrag. In twee achtereenvolgende jaren ontving een groep van respectievelijk 5 en 10 leerkrachten een training in de toepassing van directe instructie. De betrokken leerkrachten zijn zowel voor als na de training geobserveerd in hun lesgedrag en zijn vergeleken met leerkrachten die deze training niet ontvingen. De resultaten tonen aan dat leerkrachten na het volgen van de training beduidend meer tijd gaan besteden aan directe expliciete instructie in leesstrategieën en dat de kwaliteit van de instructie significant is toegenomen. De conclusie is dat een goed opgebouwde cursus voor leerkrachten ten aanzien van de instructie van begrijpend lezen belangwekkende veranderingen in de gegeven lessen teweeg kan brengen.

* Het onderzoek werd verricht in het kader van een NWO onderzoeksproject (PEDON 590-286-401).

** Hierbij willen we de betrokken leerkrachten bedanken voor hun enthousiaste medewerking aan dit onderzoek. Verder zijn we dank verschuldigd aan S. Doekes, H. Faber, S. Dölle, D. Kleijmeer, G. van Mill-van Poederooijen, M. Moons en E. de Rooij voor hun bijdrage in verschillende fasen van dit onderzoek.

1 Inleiding

Recente peilingen van de leesvaardigheid van Nederlandse leerlingen geven aanleiding om met enige verontrusting het huidige leesonderwijs onder de loep te nemen. In een vergelijkend onderzoek dat verricht is in het kader van de IEA Reading Literacy Study blijken leerlingen uit ons land verhoudingsgewijs bepaald niet tot de goede lezers te horen (De Glopper & Otter, 1992). Leerlingen uit een aantal andere landen, zoals Noorwegen, Zweden, Finland en Duitsland, presteren op tests voor begrijpend lezen duidelijk beter dan Nederlandse kinderen. Uit een recente peiling van het begrijpend lezen in groep vijf door het Cito (Sijtstra, 1993) lijkt eveneens de conclusie te kunnen worden getrokken dat de leerlingen deze vaardigheid nog maar matig tot zwak beheersen.

Over de oorzaken van deze teleurstellende gegevens tast men thans nog in het duister. De specificatie van de normen in dit type onderzoek is cruciaal en vaak discutabel. Maar als deze onderzoeken inderdaad aangeven dat er slechts povere resultaten worden bereikt door Nederlandse leerlingen, dient nader onderzoek naar de oorzaken serieus ter hand te worden genomen. Wellicht schort er iets aan het leesonderwijs in ons land? Op basisscholen - beide genoemde onderzoeken hebben betrekking op het reguliere onderwijs - wordt in de aanvangsjaren (terecht) veel tijd besteed aan technisch lezen. In latere jaren vormen vooral verhalende teksten het leesvoer. Instructie in de wijze waarop zakelijke en informatieve teksten moeten worden gelezen vindt echter nauwelijks plaats. Observaties in het basisonderwijs tonen aan, dat er sporadisch instructie wordt gegeven in het strategisch lezen van de tekst en dat leerkrachten zich gedurende lessen voor begrijpend lezen vooral richten op het voorbereiden van de vragen over een tekst en het controleren van de antwoorden (zie ook 3.1; Durkin,

1978; Weterings & Aarnoutse, 1986; Aarnoutse, 1991).

Wat geldt voor het basisonderwijs, geldt misschien ook voor het speciaal onderwijs, met name voor het lom-onderwijs waar veel leerlingen aanwezig zijn die betrekkelijk veel moeite hebben met leren lezen. Onbekend is echter in hoeverre het instructiegedrag van leerkrachten in het lom-onderwijs vergelijkbaar is met dat van collega's in het basisonderwijs tijdens lessen in begrijpend lezen. De reden van het huidige onderzoek is dan ook in de eerste plaats om hierover enige informatie in te winnen.

In recente overzichten van onderzoek naar instructie in tekstbegrip (Pearson & Fielding 1991; Reitsma & Walraven, 1991; Sliepen, 1991) wordt de behoefte aan een effectieve instructievorm onderstreept. Ook wordt in deze overzichten reeds (voornamelijk buitenlands) onderzoek met positieve effecten van deze instructievorm op het leesgedrag beschreven en zijn relevante en 'onderwijsbare' strategieën geïdentificeerd die het tekstbegrip van lezers kunnen verbeteren. Vooral het onderwijs in deze leesstrategieën volgens het directe instructiemodel lijkt vrucht af te werpen. Maar deze aanpak is thans blijkbaar nog nauwelijks geïmplementeerd in het reguliere onderwijs (cf. Weterings & Aarnoutse, 1986). Is dit in het lom-onderwijs wel het geval?

Gezien de populatie van leerlingen en het speciale karakter van het onderwijs - veelal een intensieve en gespecialiseerde begeleiding van leerlingen met leerproblemen - zou men kunnen verwachten dat in het lom veel, of althans meer, expliciete en directe instructie gegeven wordt in begrijpend lezen. Op lom-scholen bevinden zich relatief veel zwakke lezers die vooral moeite hebben met het verwerven van de technische leesvaardigheid. Het leesonderwijs is dikwijls speciaal gericht op het verbeteren van het technisch lezen. Maar ook zwakke lezers krijgen, naarmate ze ouder worden en in hogere groepen onderwijs ontvangen, steeds meer schriftelijke informatie te verwerken op school. In allerlei zaakvakken wordt hun leesvaardigheid steeds meer op de proef gesteld. De toename in de hoeveelheid te verwerken tekst gaat echter niet altijd of niet automatisch gepaard met een toename in technische leesvaardigheid en tekstbegrip. Specifieke instructie en begeleiding is dus nodig voor deze leerlingen. Maar wordt deze ook gegeven?

In deze bijdrage wordt verslag gedaan van twee studies, die zijn gericht op het beschrijven en veranderen van instructiegedrag van leerkrachten tijdens lessen in begrijpend lezen in het lom-onderwijs. De eerste studie is een kleinschalig onderzoek, waarin de methode van onderzoek wordt beproefd en een schatting van de tijdsinvestering door middel van observaties plaatsvindt. Het tweede onderzoek is een uitgebreide en verbeterde opzet, waarin tevens het effect van een training van leerkrachten wordt geëvalueerd.

2 Instructie in begrijpend lezen

Begrijpend lezen is het proces waarin een lezer de informatie die in een tekst in- of expliciet aanwezig is verwerkt en integreert met zijn kennis omtrent de context en kennis van de omringende wereld. Een goed begrip van de tekst vereist een veelheid aan interacties met deze tekst. Geschreven woorden worden visueel waargenomen en geïdentificeerd. Betekenis wordt toegekend aan zinsdelen en/of zinnen. Bij deze betekenistoekenning speelt reeds bestaande kennis een rol. Een integratie van de verschillende betekeniseenheden uit de tekst dient plaats te vinden om tot een coherent beeld te komen van hetgeen in een tekst beschreven staat. De nuancering en diepte van de betekenisverlening is o.a. afhankelijk van het doel waarmee men leest, hoeveel kennis er al is over het betreffende onderwerp en hoeveel ervaring men heeft met de structuur van dergelijke teksten. Het is buitengewoon lastig om dit interactieve proces van begrijpend lezen, waarin zoveel elementen een rol spelen, in detail te beschrijven. Maar het mag wel duidelijk zijn dat begrijpend lezen meer is dan het louter decoderen van tekst.

Voor begrijpend lezen worden van de lezer actieve (meta)cognitieve processen verlangd. Goede lezers maken efficiënt gebruik van decodeerprocessen en besteden hieraan niet meer tijd en energie dan nodig is. De aandacht gaat vooral uit naar een variëteit aan strategieën die van belang zijn voor het bereiken van een begrip van de tekst en eventueel het onthouden van hetgeen is gelezen. Ze schakelen voor-

kennis in voor zover nodig of gewenst, benutten de context om nieuwe of ambigue woorden of zinsdelen te interpreteren, maken gebruik van kennis omtrent tekststructuren (alineaindeling, tussentitels, e.d.) en bewaken voortdurend of de interpretatie van de tekst een samenhangend geheel is dat overeenkomt met (eerder gewekte) verwachtingen of het leesdoel. Metacognitieve activiteit bestaat hierin dat lezers doelgericht een tekst gaan lezen, het begrip tijdens het lezen of vlak na een passage controleren of evalueren en bij gebleken onbegrip passende 'reparatie'-strategieën inschakelen. Goede lezers geven er meestal blijk van deze metacognitieve vaardigheden en strategieën toe te passen tijdens begrijpend lezen. Zwakke lezers daarentegen richten zich vaak voornamelijk op het decoderen van losse woorden, passen het leesgedrag niet aan aan het leesdoel en het soort tekst en kijken zelden terug of vooruit in de tekst om het leesbegrip te verhogen. Deze afwezigheid van dergelijke leesstrategieën is kenmerkend voor zowel jonge onervaren lezers als oudere zwakke lezers. Oudere zwakke lezers in het lom-onderwijs vertonen bovendien vaak motivatieproblemen, omdat ze in de loop der tijd lage succesverwachtingen hebben gekregen en snel opgeven zodra er leesproblemen optreden (Paris, Wasik & Turner, 1991).

Over de *vorm* waarin instructie in begrijpend lezen het meest optimaal kan worden gegeven is door onderzoek reeds het een en ander bekend. Op grond van recente resultaten lijkt een advies aan leerkrachten voor de hand te liggen: als leerkrachten het tekstbegrip van (zwakke) lezers willen verbeteren, is het raadzaam het flexibel gebruik van leesstrategieën expliciet en direct te onderwijzen (Pearson & Fielding, 1991; Reitsma & Walraven, 1991). Het model van directe instructie lijkt bijvoorbeeld uiterst effectief. Het klassikaal gerichte onderwijsgedrag van de leerkracht wordt in deze aanpak gekenmerkt door een directe en systematische manier van lesgeven. De leerkracht legt daarbij expliciet uit welke deelvaardigheid of strategie er geleerd gaat worden, hoe, wanneer en waarom deze in een tekst kan worden toegepast. Voorafgaand aan deze instructie is er een introductiefase en erop volgend zijn er een oefenings- en verwerkingsfase. De leerstof wordt in kleine stappen aangeboden, er wordt expliciet uitgelegd waar het om gaat, er wordt door de leerkracht voorgedaan hoe het gaat (modelling), er worden voldoende voorbeelden gegeven en bij elke stap wordt intensief geoefend. De oefeningen worden verder nauwgezet begeleid en becommentarieerd (feedback). Gedurende het oefenproces wordt geleidelijk aangestuurd op een zelfstandige verwerking van de leerstof en vermindert de begeleiding (Roehler & Duffy, 1991). In het op kleine groepen gerichte model van 'wederkerig onderwijzen' (reciprocal teaching) kan deze aanpak nog intensiever worden gevolgd. De leerkracht heeft aanvankelijk de leiding in een discussie over de tekst. Hij stelt vragen over passages of woorden in de tekst, maakt (tussentijdse) samenvattingen en voorspellingen. De bedoeling is dat de leerlingen dit modelgedrag - hardop denkend de tekst trachten te vatten - zullen opvolgen. Zij krijgen daarom beurtelings ook de rol van de discussieleider. Aanvankelijk ondersteunt de leerkracht deze activiteit, maar hij trekt zich geleidelijk terug. Op deze wijze wordt bevorderd dat de leerlingen zowel ter zake relevante leesstrategieën direct toepassen als een gelijktijdige controle op het begrip van de tekst (monitoring) uitvoeren. Onderzoek laat zien dat deze onderwijsstrategie goede resultaten oplevert (o.a. Pennings, De Jong & Laudy, 1989; Palincsar & Brown, 1989).

De *inhoud* van de instructie in begrijpend lezen heeft meestal betrekking op deelvaardigheden en leesstrategieën, die in onderzoek reeds van belang zijn gevonden voor een goed begrip van een tekst (Pearson, Roehler, Dole & Duffy, 1991; Reitsma & Walraven, 1991; Sliepen, 1991). Het betreft o.a. instructie in het zichzelf vragen stellen tijdens het lezen, het actief gebruik leren maken van voorkennis tijdens het lezen, het rekening leren houden met de structuur van een tekst (titels, alinea's, interpunctie, e.d.), een voortdurende controle op semantische samenhang en consistentie, het onderscheid tussen belangrijke en onbelangrijke zaken, het leren trekken van (tussentijdse) conclusies, het maken van inferenties tijdens en na het lezen en het leren samenvatten. Het is van belang deze leesstrategieën niet als geïsoleerde vaardigheden te oefenen, maar deze altijd in de normale, zinvolle context van de lezing van een tekst aan de orde te stellen.

3 De praktijk van begrijpend leesonderwijs

3.1 De praktijk in het basisonderwijs

Hoewel, zoals bovenbeschreven, in grote lijnen bekend is hoe instructie in begrijpend lezen er uit zou kunnen of moeten zien, kan de vraag rijzen of dit in de praktijk ook zo gerealiseerd wordt. Eerder onderzoek in het reguliere basisonderwijs levert een krachtig ontkennend antwoord. Nadat Durkin (1984) reeds in de Verenigde Staten had aangetoond dat in het basisonderwijs bijzonder weinig echte instructie in begrijpend lezen werd gegeven, is door Weterings en Aarnoutse (1986) een vergelijkbaar onderzoek in Nederland verricht.

Weterings en Aarnoutse (1986) observeerden in totaal bijna 100 lessen bij 12 leerkrachten in groep 6 van de basisschool. Met behulp van een categorieënsysteem dat bestond uit 36 observatiecategorieën werden alle leerkrachtactiviteiten tijdens lessen in begrijpend lezen geclassificeerd. Bovendien werd nagegaan hoeveel tijd aan deze activiteiten besteed werd door op de minuut nauwkeurig aan te geven welke leerkrachtactiviteit geobserveerd werd gedurende de les. Uit de resultaten bleek dat de leerkrachten ongeveer 25% van de tijd besteden aan het beoordelen van het begrijpend lezen, dat wil zeggen, luisteren naar antwoorden van leerlingen. Andere veel tijd vergende activiteiten waren o.a.: helpen met opdrachten, stellen van vragen, uitleggen wat de bedoeling van een vraag is en luisteren naar (hardop) leesbeurten. Als alle activiteiten die met begrijpend lezen te maken hebben samen worden genomen, blijkt hier ongeveer de helft van de lestijd aan te worden besteed. Maar de hoofdbestanddelen van deze activiteiten rond begrijpend lezen zijn het controleren van en het helpen bij opdrachten. Aan werkelijke instructie - de leerkracht doet iets om leerlingen te helpen de betekenis te begrijpen van meer dan één woord, geeft uitleg met betrekking tot het beantwoorden van vragen, geeft onderwijs in studievaardigheden, of geeft uitleg over woordbetekenissen - wordt slechts 4% van de lestijd besteed.

In een vervolgonderzoek in groep 7 van de basisschool worden deze bevindingen in feite gerepliceerd (Aarnoutse, 1991). De 22 geobserveerde leerkrachten zijn gedurende meer dan 25% van de tijd bezig met allerlei procedurele zaken en blijken bijna 18% van de tijd niet bij de les betrokken (ze kijken dan ander werk na, e.d.). Verder wijzen de resultaten uit dat ze ongeveer 17% van de tijd klassikale instructie geven, maar slechts een zeer klein deel hiervan wordt besteed aan specifieke instructie in begrijpend leesstrategieën. Aan de hand van observaties van de leerlingen blijken de leerlingen namelijk slechts 0.1% van de tijd instructie te ontvangen in vaardigheden of strategieën voor begrijpend lezen.

Uit deze onderzoeksbevindingen kan dus worden geconcludeerd dat echte instructie in begrijpend lezen grotendeels ontbreekt in de onderzochte groepen. Naar de oorzaak is niet gespeurd, maar een voor de hand liggende verklaring is dat de gehanteerde leesmethoden geen richtlijnen geven voor deze wijze van instructie. Zoals bovenbeschreven zijn de inzichten op het terrein van de instructie in begrijpend leesstrategieën tamelijk recent. De kennis op dit gebied is ook nog beperkt en is voor zover beschikbaar zeker niet volledig doorgevoerd in de methoden voor voortgezet lezen. In de scholen werden nog methoden gebruikt die niet of nauwelijks aanwijzingen bevatten voor specifieke instructie. In deze situatie kan inmiddels overigens verandering zijn gekomen, omdat nieuwe methoden op de markt zijn gekomen waarin deze nieuwe aanpak met specifieke aandacht voor daadwerkelijke instructie wel is opgenomen.

3.2 De praktijk in het lom-onderwijs

De resultaten van onderzoek naar de praktijk van instructie in begrijpend lezen in het basisonderwijs kan niet zonder meer geldig worden verklaard voor de situatie in andere typen onderwijs. In het lom-onderwijs wordt doorgaans in kleinere groepen gewerkt en wordt vermoedelijk meer individueel gericht onderwijs gegeven. Veel leerlingen op lom-scholen hebben vanwege een frequente combinatie van leer- en gedragsmoeilijkheden bovendien een duidelijke en expliciete instructie nodig. Op grond hiervan zou men kunnen verwachten dat een observatiestudie in het lom-onderwijs aan het licht zou brengen dat hier meer directe instructie in begrijpend lezen wordt verstrekt.

Onderzoek met behulp van vragenlijsten geeft aanleiding tot de veronderstelling dat in het lom-onderwijs daadwerkelijk meer aan-

dacht wordt besteed aan begrijpend lezen dan in het basisonderwijs, niet zozeer in de instructievorm, maar wel in de tijdbesteding. In totaal wordt meer tijd besteed aan begrijpend lezen dan in het basisonderwijs. In een voorstudie periodieke peiling van het onderwijsniveau in het speciaal onderwijs (Otter, Oostdam & Vergeer, 1988) kwam van 106 lom-scholen als antwoord op een toegezonden vragenlijst dat in de eindgroepen per week gemiddeld 7,5 uur aan moedertaalonderwijs wordt besteed. Deze totale tijd voor taal bedraagt twee uur meer dan in de achtste groep van het basisonderwijs. Bovendien wordt in het lom-onderwijs van de totale taaltijd ongeveer 50% aan het onderdeel lezen besteed, terwijl dit in het basisonderwijs slechts 20% is. De tijd die aan lessen speciaal gericht op begrijpend lezen wordt besteed, is gemiddeld ongeveer 50 minuten per week. Op grond van vragenlijstgegevens kan uiteraard slechts een zeer globaal beeld worden geschetst, maar het pleit zeker niet tégen de veronderstelling dat in het lom-onderwijs meer aandacht aan instructie in begrijpend lezen wordt besteed.

Anderzijds is te verwachten dat in het lom-onderwijs eigenlijk eenzelfde situatie aan te treffen zal zijn als in het basisonderwijs. Omdat er vermoedelijk een nauwe samenhang bestaat tussen de aanbevolen instructieprocedures in de leesmethoden en de instructies die leerkrachten tijdens de les ook werkelijk geven, zal bij gebrek aan geschikte leesmethoden ook in het lom-onderwijs (nog) betrekkelijk weinig specifieke instructie te bespeuren zijn. Het lom-onderwijs maakt vaak noodgedwongen gebruik van leesmethoden voor het basisonderwijs, omdat er slechts in beperkte mate speciaal voor lom-onderwijs geschikte methoden worden uitgebracht. Zoals eerder vermeld, wordt recentelijk weliswaar in enkele nieuwe leesmethoden aandacht besteed aan instructie in leesstrategieën, maar deze zijn veelal geschreven voor de bovenbouw van het basisonderwijs of het voortgezet onderwijs en daarmee leestechnisch over het algemeen te moeilijk voor leerlingen in het lom-onderwijs. Kortom, als er leesmethoden gebruikt worden waarin niet gerept wordt over directe instructie in leesstrategieën, is nauwelijks te verwachten dat leerkrachten hierin wel onderwijs geven, tenzij zij wellicht in de laatste jaren een nascholing op dit terrein hebben gevolgd (Barr & Sadow, 1989; Pearson & Fielding, 1991).

3.3 Observatie van de lespraktijk

Voorafgaand is al gerefereerd aan onderzoeken waarin de praktijk van leesinstructie is onderzocht. Het gebruik van een vragenlijst geeft slechts een algemene indicatie van het onderwijsgedrag. Zowel het globale karakter van een vragenlijst als de niet of nauwelijks te controleren mogelijkheid tot antwoordtendenties maakt een vragenlijst weinig geschikt om precies na te gaan welke activiteiten leerkrachten tijdens hun lessen ontplooien. Observaties van actueel lesgedrag zijn noodzakelijk om een nauwkeuriger beeld te verkrijgen. Er zijn echter diverse mogelijkheden voor observatiestudies en de verschillen in methodologie zijn van belang voor de vragen die men kan beantwoorden.

Weterings en Aarnoutse (1986) classificeerden de leerkrachtactiviteiten in 34 inhoudelijke observatiecategorieën met 1 minuut als basiseenheid van observatie. Deze wijze van observatie betekent dat activiteiten, die een kortere tijd in beslag namen, niet werden geregistreerd. In een latere studie werd door Aarnoutse (1991) zowel een leerling als de leerkracht (gelijktijdig) gedurende 20 seconden geobserveerd en tijdens de volgende 10 seconden werd er gescoord; daarna begon een nieuwe observatieperiode (met dezelfde leerkracht, maar een andere leerling). Gescoord werd hetgeen als hoofdactiviteit tijdens elke 20 seconden kon worden waargenomen. De kortere periode voor observatie noopte de onderzoekers echter tot een drastische reductie van het aantal categorieën voor de scoring van leerkrachtgedrag. In plaats van 34 categorieën werden nu slechts 5 inhoudelijke gedragscategorieën gebruikt. Een belangrijke keuze die bij observaties gemaakt moet worden, namelijk de tijdseenheid van observeren, blijkt hier onmiddellijk consequenties te hebben voor het aantal te onderscheiden gedragscategorieën. En in het algemeen is het omgekeerde ook waar: een genuanceerde beoordeling vergt nu eenmaal veel tijd en vereist daarom langere intervallen voor observatie. Een keuze voor korte intervallen en een nauwkeurige observatie van kleine gedragseenheden maakt het mogelijk relatief objectief gegevens te verzamelen, maar dwingt

tot een interpretatie van de gegevens achteraf. En het blijkt veelal lastig om achteraf op grond van de gegevens omtrent moleculaire gedragseenheden tot een algemene en zinvolle beschrijving te komen van het (molaire) gedrag.

Als de vraag gesteld wordt of leerkrachten instructie in begrijpend lezen geven, dan is één manier van onderzoek om met een categorieënsysteem voor observatie op een kwantitatieve wijze (time-sampling) de verschillende leerkrachtactiviteiten in kaart te brengen. De uitkomsten geven een benadering van de tijd die een leerkracht aan bepaalde activiteiten besteedt of de frequentie waarin een bepaalde handeling voorkomt. Uit het patroon van de tijdbesteding aan verschillende activiteiten of de frequentieverdeling over verschillende categorieën kan een indicatie worden verkregen van de kwaliteit van de totale les. Echter omtrent de kwaliteit van de instructie zelf wordt hiermee weinig bekend. Is een (breedsprakige) leerkracht die veel tijd besteedt aan instructie ook daadwerkelijk beter dan een collega die in een paar (kernachtige) woorden een strategie kan uitleggen? Deze kwalitatieve toevoegingen tussen haakjes kunnen met het categorieënsysteem onvoldoende worden geregistreerd. Voor een dergelijke kwalitatieve beoordeling van de instructie is een andere beoordelingswijze noodzakelijk.

In het eerste hier gerapporteerde onderzoek wordt getracht de complexiteit van een instructiesituatie weer te geven door op twee manieren gegevens te verzamelen. Naast een kwantitatieve manier van observatie van de verschillende leerkrachtactiviteiten wordt ook een waarderingschaal gehanteerd. Het categorieënsysteem voor kwantitatieve observatie is samengesteld naar aanleiding van o.a. de observatieschalen van Weterings en Aarnoutse (1986), Barr en Sadow (1989) en Aarnoutse (1991). De waarderingschaal is een aanpassing van een schaal die elders reeds gebruikt is en waardevol is gebleken (Duffy, Roehler, Sivan, Rackliffe, Book, Meloth, Vavrus, Wesselman, Putnam & Bassiri, 1987). Duffy en zijn collega's observeerden lessen in begrijpend lezen bij leerkrachten in het basisonderwijs en beoordeelden na afloop van een les op vijfpuntschalen de mate waarin strategieën expliciet aan de leerlingen werden overgedragen. De vraag kan nu gesteld worden in welke mate de beide manieren van gegevensverzameling - systematische observaties in korte tijdsintervallen en een kwalitatieve waardering van de gehele les - overeenkomstige resultaten produceren. Als beide een indicatie zijn van de kwaliteit van lessen - en dat is de assumptie van beide methoden van observatie —, dan is een redelijke overeenkomst te verwachten. Indien echter het patroon van tijdbesteding binnen een les betrekkelijk onafhankelijk is van de kwalitatieve waardering van de les, dan zullen er (grote) verschillen optreden.

4 Onderzoek I

4.1 Probleemstelling

In dit onderzoek wordt getracht in eerste instantie antwoord te geven op de volgende vragen: 1) Welke leerkrachtactiviteiten kunnen worden waargenomen tijdens lessen in begrijpend lezen in eindgroepen van het lom-onderwijs, 2) in hoeverre is er bij deze activiteiten sprake van directe instructie in leesstrategieën, en 3) in hoeverre komen beschrijvingen op grond van systematische observatie gedurende kleine tijdsintervallen overeen met een meer algemene kwaliteitswaardering?

Omdat naar verwachting een instructievorm en -inhoud die voldoen aan de specificaties van het directe instructiemodel nauwelijks worden aangetroffen, wordt als derde onderzoeksvraag gesteld: 4) in hoeverre het mogelijk is de instructieactiviteiten van deze leerkrachten te veranderen in de gewenste richting door hen te trainen in het geven van directe instructie in leesstrategieën in de dagelijkse onderwijspraktijk.

4.2 Onderzoeksgroep

Voor dit onderzoek zijn uitsluitend leerkrachten benaderd uit de bovenbouw van het lom-onderwijs, die les geven aan 10 tot 13-jarige leerlingen met een technisch leesniveau vergelijkbaar met groep 4 en 5 basisschool (A VI ≥4 - ≤9). Aan hen is gevraagd deel te nemen aan het onderzoek waarbij als doelen waren vermeld: a) een inventarisatie van het onderwijs in begrijpend lezen aan oudere zwakke lezers door middel van (video)observaties in de klas, en b) het beproeven van een nieuwe aanpak voor instructie in begrijpend lezen door deelname van

enkele leerkrachten aan een training over begrijpend leesinstructie voor zwakke lezers en het geven van bijbehorende proeflessen in de klas. Er werden 20 scholen benaderd. Vanwege tijdgebrek en allerlei organisatorische redenen bleken uiteindelijk slechts 5 leerkrachten van bovenbouwgroepen te kunnen participeren. Wel was er een flink aantal belangstellende leerkrachten van jongere leerlingen, maar omdat het onderzoeksproject zich richt op de bovenbouw van het lom kwamen zij niet in aanmerking.

De onderzoeksgroep waarover hier wordt gerapporteerd bestaat uit twee vrouwen en drie mannen, die les geven op vier verschillende lom-scholen. Deze vijf leerkrachten hebben drie eindgroepen en twee schakelklassen vso-lom onder hun hoede en de groepen bestaan gemiddeld uit 17 leerlingen. De leerkrachten hebben gemiddeld 13.5 jaar onderwijservaring en hiervan heeft gemiddeld 6.5 jaar betrekking op ervaring in het lom-onderwijs. Vier van de vijf leerkrachten geven ongeveer één uur in de week les in begrijpend lezen en hanteren alle vier een andere methode. Een vijfde leerkracht heeft geen speciale methode voor begrijpend lezen en trekt hiervoor ook niet een apart lesuur uit. Eén leerkracht geeft les aan een betrekkelijk homogene leesgroep, omdat er door de hele school niveaugroepen voor begrijpend lezen zijn samengesteld. Bij de overige vier leerkrachten treffen we heterogene leesgroepen aan, met zowel zwakke als goede technische lezers.

4.3 Procedure

In totaal zijn 38 lessen in begrijpend lezen voor nadere codering op videoband vastgelegd. De lessen van vijf leerkrachten zijn voorafgaand aan de cursus vier maal en verspreid over ruim twee maanden geobserveerd (vanaf video). De totale observatietijd tijdens deze eerste 20 lessen bedraagt in totaal 766 minuten en varieert per les van 29 tot 51 minuten.

Voorafgaand aan de observaties is in elke klas de videocamera geïntroduceerd en is een proefles opgenomen. Deze videobeelden werden aan de leerkracht en leerlingen getoond. Alle observaties vonden per leerkracht steeds op dezelfde dag in de week en op hetzelfde tijdstip plaats. De camera volgde het gedrag van de leerkracht steeds vanuit eenzelfde positie in de klas. De leerkracht droeg een dasspeldmicrofoon opdat ook alle verbale uitingen verstaanbaar geregistreerd konden worden. De leerlingen kwamen alleen in beeld tijdens een interactie met de leerkracht.

Elke opgenomen les is achteraf eerst met het categorieënsysteem (zie 4.4 voor een beschrijving van de categorieën) door twee vooraf getrainde observatoren onafhankelijk kwantitatief beoordeeld: het gedrag van de leerkracht op of rond elke 10e seconde werd gecodeerd. De codering en de video-opname startte op het moment dat de leerkracht aangaf met de les te beginnen. Als op het moment van observatie een adempauze viel werd het gedrag er vlak aan voorafgaand gecodeerd. Voor een juiste interpretatie werd de band regelmatig teruggespoeld of werd het gedrag van de volgende paar seconden erbij betrokken. De videobeelden zijn nadien nogmaals onafhankelijk bekeken en nu werden uitsluitend instructie-activiteiten genoteerd die betrekking hadden op de vragen van de waarderingschaal. De verbale uitingen van de leerkracht werden zoveel mogelijk letterlijk, in steekwoorden opgeschreven. Na afloop van de les werd de kwaliteit van de activiteiten beoordeeld door toekenning van een cijfer. De afwezigheid van het criterium werd aangegeven met een 0 en het perfecte voorbeeld met een 4.

Vervolgens is begonnen met vijf wekelijkse trainingsbijeenkomsten van ca. 2.5 uur. De vijf leerkrachten hebben tijdens en na de training gelegenheid gehad om zelfstandig met directe instructie in strategisch lezen te oefenen. Naast voorbeeldlessen is door de leerkrachten aanvullend lesmateriaal gezocht voor de resterende maanden tot de zomervakantie. Daarbij kon worden uitgegaan van de teksten uit de in gebruik zijnde methode voor begrijpend lezen.

Na de cursus zijn verspreid over ruim twee maanden nogmaals enkele lessen op video vastgelegd en volgens dezelfde procedure als boven beschreven beoordeeld. Bij drie leerkrachten is vier maal en bij twee drie maal geobserveerd. Tijdens deze 18 lessen van de tweede observatieperiode is in totaal 680 minuten geobserveerd; de lessen variëren van 19 tot 55 minuten.

4.4 Observatieinstrumenten

Voor het registreren van de activiteiten van

leerkrachten tijdens lessen in begrijpend lezen is gebruik gemaakt van twee verschillende observatieschalen: een categorieënsysteem voor observaties per tijdsinterval en een algemene schaal voor het waarderen van de kwaliteit.

Omdat we vermoeden dat leerkrachten nauwelijks tijd besteden aan directe instructie in leesstrategieën en we niet precies weten waaraan zij hun tijd dan wel besteden, is het categorieënsysteem ruim opgezet. Niet alleen instructie, maar ook andere activiteiten kunnen worden ondergebracht. Voor de observaties per tijdseenheid worden 12 categorieën onderscheiden (zie ook Tabel 1). In het observatiesysteem wordt op elke 10e seconde geregistreerd of een leerkracht tijd besteedt aan één van de vier algemene categorieën die niet specifiek op begrijpend lezen gericht zijn of aan één van de acht categorieën die specifiek zijn voor instructie in begrijpend lezen. De vier algemene categorieën zijn: - niet lesgericht (bijv. ander werk nakijken), - algemene procedures (bijv. uitdelen of verzamelen van leermiddelen; orde handhaving), - technisch lezen (bijv. geven van beurt voor hardop lezen; hulp bij of correctie van decoderen), - spelling/syntaxis (bijv. waardering van schrijfwijze van antwoorden). De acht categorieën die specifiek zijn voor instructie in begrijpend lezen zijn: - voorbereiden op de te lezen tekst (bijv. opwekken van achtergrondinformatie, formuleren van leesdoel); - voorspellen van de inhoud van (delen van) de tekst (bijv. genereren van verwachtingen naar aanleiding van de titel); - conclusies trekken over de inhoud en de juistheid van de gedane voorspelling; - instructie over specifieke tekstkenmerken (zoals bijv. tekststructuur, interpunctie, referentiële relaties); - instructie over studievaardigheden (zoals bijv. omgaan met een woordenboek, schema's maken, onderstrepen van tekstdelen); en - luisteren. Met luisteren wordt hier bedoeld, het luisteren naar leerlingen die vragen stellen of beantwoorden over de tekst, die verwachtingen formuleren en toetsen, of een andere oplossingsstrategie verwoorden. Er wordt niet onder verstaan, het luisteren naar andere inhoudelijke activiteiten van leerlingen, zoals bijvoorbeeld niet tekstgerichte discussies (dit valt onder algemene procedures), of het verklanken van woorden tijdens het hardop lezen van een tekst of opdracht (dit wordt ondergebracht bij de categorie technisch lezen). Daarnaast waren er nog twee categorieën die te maken hebben met de leesprestaties: - resultaatgericht begeleiden (bijv. het controleren van de oefeningen); en - leerprocesgericht waarderen (bijv. het prijzen van een goede aanpak, een aanmoediging om door te gaan met de gekozen oplossingstrategie). Hoewel de gerichtheid van de leerkracht ook werd gecodeerd (individeel, subgroep of klassikaal) wordt deze nadere opsplitsing hier niet gerapporteerd (voor zover instructie plaatsvond was deze klassikaal). Voor het materiaal in dit onderzoek is de inter- en intracodeursbetrouwbaarheid bepaald door Cohen's Kappa (Siegel & Castellan, 1988) te berekenen: de intra-beoordelaarsbetrouwbaarheid was .82 en de interbeoordelaarsbetrouwbaarheid was .78. In het huidige onderzoek zijn bij afwijkingen tussen twee observatoren de waarnemingen gemiddeld voordat de gegevens verder werden verwerkt.

Elke les werd tevens beoordeeld op een ordinale schaal (zie appendix). Deze schaal is een vertaling en aanpassing van de 'Rating categories for explicitness of teacher explanations' (Duffy e.a., 1987) en kan worden gebruikt om te beoordelen hoe expliciet een leerkracht is in zijn uitleg over het hoe, wanneer en waarom van bepaalde strategische aanpak van een tekst. De woordelijke uitleg van de leerkracht staat centraal (Duffy e.a., 1987). Deze waarderingschaal bestaat uit elf onderwerpen verdeeld over drie categorieën: (1) de presentatie van de informatie over een strategie, (2) de lesopbouw waarbinnen deze informatie wordt geboden en (3) de samenhang tussen de lessen en de lesonderdelen. Het eerste deel - informatie over de strategie - gaat over hetgeen de leerkracht tegen de leerling(en) zegt over de te leren vaardigheid of aanpak, de bruikbaarheid ervan, wanneer je hiervoor kunt kiezen, welke (mentale) processen hierbij een rol spelen, e.d. Het tweede deel - de lesopbouw - is gericht op de (klassikale) introductie van de les, het (stapsgewijs) voordoen van de strategie of vaardigheid, het geleidelijk verminderen van de hulp, afronding van een les, e.d. In dit tweede deel zijn componenten van het directe instructiemodel te herkennen: de introductie, presentatie, inoefening, verwerking en evaluatie. Het derde deel heeft betrekking op de samenhang tussen de lesonderdelen en/of de verschillende lessen, of

er een logische sequentie in de uitleg zit zonder inconsistenties of onderbrekingen, relevante verwijzingen naar vorige lessen, e.d. De inter- en intrabeoordelaarsbetrouwbaarheid (Goodman-Kruskal's Gamma; zie Siegel & Castellan, 1988) van deze schaal bleek respectievelijk .93 en .95. Ook hier zijn bij eventuele afwijkingen tussen twee observatoren de scores voor verdere verwerking gemiddeld.

4.5 Training van leerkrachten

Om de informatie over directe instructie in leesstrategieën over te dragen is gekozen voor een trainingsopzet die zo dicht mogelijk aansluit bij het model van Joyce en Showers (in Tillema & De Jong, 1991). Het genoemde model bevat juist die componenten, die concrete en zichtbare gedragsveranderingen bij volwassenen tot stand kunnen brengen, namelijk:

a *Diagnose.* Twee maanden voor aanvang van de training, aan het begin van de observaties in de klas, hebben de leerkrachten een vragenlijst ingevuld over het begrijpend leesonderwijs in hun klas. Daarmee hebben we enig inzicht kunnen krijgen in de begin-situatie, voorkennis, verwachtingen en behoeften. Zo werden bijvoorbeeld een gebrekkige woordenschat en schriftelijke verwerking van de tekst genoemd als meest gesignaleerde problemen bij de leerlingen tijdens begrijpend lezen. Deze informatie is tijdens de eerste bijeenkomst ook aan de deelnemers teruggerapporteerd, samen met informatie over doel en inhoud van de training.

b *Informatie, demonstratie / modelling.* In tegenstelling tot de voorgaande component die met name ter voorbereiding op en tijdens de eerste bijeenkomst aandacht krijgt, worden deze en volgende componenten elke bijeenkomst als volgt aan de orde gesteld. Het doel per bijeenkomst en het belang van de te behandelen onderwerpen worden duidelijk gemaakt door verwijzing naar resultaten van wetenschappelijk onderzoek op dit terrein. Door het presenteren van informatie over het onderwerp wordt kennis over het onderwerp bij cursisten geactiveerd en aangevuld of bijgesteld. De twee hoofdonderwerpen: het directe instructiemodel en een strategische aanpak van teksten, worden expliciet en stap voor stap uitgelegd en met behulp van voorbeeldteksten voor zowel volwassenen als leerlingen, gedemonstreerd. De toepassing en bruikbaarheid in de klaspraktijk wordt gedemonstreerd met behulp van videobeelden met modelgedrag (Boonman, Kok & De Haan, 1988; Walraven, 1990).

c *Praktijkoefening en terugkoppeling.* Aan deze component van de training wordt op twee manieren aandacht besteed. Ten eerste wordt de cursisten gevraagd, om bij videobeelden van eigen lesgedrag na te gaan in hoeverre de zojuist behandelde aspecten in het lesgedrag aanwezig zijn en om concrete adviezen te geven voor eventuele bijstellingen. Hiervoor worden (vooraf geselecteerde) fragmenten gebruikt van de op video vastgelegde lesobservaties. Voor aanvang van de training hebben we de leerkrachten hiervoor uitdrukkelijk om toestemming gevraagd, omdat het weliswaar een effectief, maar tegelijkertijd zeer indringend en confronterend leermiddel is.
Daarnaast ontvangen de leerkrachten van elk behandeld onderwerp een voorbeeldles opgebouwd volgens het directe instructiemodel, waarin de instructie volledig staat uitgeschreven en is er tijdens de vijf trainingsbijeenkomsten tijd gereserveerd om samen met een collega zelf een les te schrijven, analoog aan het voorbeeld. Deze les wordt dan aansluitend besproken. Beide lessen kunnen in de praktijk worden uitgevoerd en op een volgende trainingsbijeenkomst worden geëvalueerd. Om klassikale (in plaats van groeps- of individueel gerichte) instructie mogelijk te maken, zijn er informatieve teksten geselecteerd en geschreven die ook voor zwakke lezers makkelijk leesbaar zijn (AVI-4 niveau). We gaan ervan uit dat denkstrategieën klassikaal kunnen worden aangeleerd en geoefend; voor de uiteindelijke verwerking en zelfstandig toepassing op andere leestaken kunnen eventueel teksten worden gekozen die differentiëren naar leesniveaus. De onderwerpen (strategieën) die achtereenvolgens per les behandeld worden zijn: soorten teksten en bijbehorende manieren om te lezen, het activeren van kennis over het onderwerp, het voorspellen van de inhoud en het toetsen en eventueel bijstellen van die

verwachtingen, problemen tijdens controle van het begrip zoals aanpak van onbekende woorden en het doorzien van relatie-aanduidingen in de tekst zoals oorzaak-gevolg, middel-doel en antecedent en anafoor.

d *Coaching.* Ondersteuning en coaching is een essentieel deel van de training dat node gemist kan worden als transfer of implementatie van het geleerde wordt beoogd (Snippe, 1991), maar deze component komt in het huidige onderzoek alleen tijdens de trainingsbijeenkomsten aan de orde. De leerkracht kan tijdens de bijeenkomsten problemen tijdens de uitvoering in de klas bespreken. In die zin is er sprake van ondersteuning, maar deze vindt niet individueel plaats op de werkplek. Voor een individuele begeleiding is niet gekozen om het aantal oncontroleerbare factoren in het onderzoek niet onnodig te vergroten. Door alleen tijdens de bijeenkomsten te evalueren en te adviseren, konden we ervoor zorgen dat aan elke leerkracht grotendeels dezelfde informatie verstrekt werd. Deze aanpak, samen met de volledig uitgeschreven lessen en de videobeelden moesten het ontbreken van verdere coaching in de eigen klas zoveel mogelijk ondervangen. Omdat bij elke leerkracht na de training nog enkele malen een les op video zou worden vastgelegd, was te verwachten dat de leerkrachten zeker enige tijd zouden proberen om de nieuwe kennis zo goed mogelijk in praktijk te brengen.

Tot nu toe zijn alleen de hoofdonderwerpen 'directe instructie' en 'strategische aanpak' genoemd. Meer concreet kan gesteld worden, dat het directe instructie-model eerst in totaliteit is gepresenteerd als instructie-vorm en vervolgens is toegepast op elke nieuwe instructie-inhoud: de leesstrategie(ën). Vooral aan twee componenten van het directe instructiemodel is speciale aandacht besteed: het hardop verwoorden van de achtereenvolgende denkstappen bij het toepassen van een strategie (modelling) en het geven van gerichte terugkoppeling op de wijze waarop de leerling een strategie heeft uitgevoerd (feedback).

In de korte training kunnen uiteraard niet alle leesstrategieën even uitvoerig aan de orde komen. In navolging van o.a. Paris, Cross en Lipson (1984) en Paris en Oka (1986) zijn twee groepen strategieën onderscheiden: strategieën direct gericht op het construeren van de betekenis van een tekst en strategieën die meer algemeen gericht zijn op het bewaken en sturen van het lees- en begripsproces. Als representant van de eerste groep is gekozen voor de aanpak van onbekende woorden en causale relaties met in het bijzonder de relatie tussen antecedent en anafoor. De laatste groep is weergegeven in de vorm van vragen die de leerlingen zichzelf kunnen stellen, voor, tijdens en na het lezen. Bijvoorbeeld voorafgaand aan het lezen: Wat is dit voor een tekst en hoe ga ik dit lezen? Wat weet ik er al van en waar zal het over gaan? Tijdens het lezen: Klopt het wat ik dacht? Het wordt moeilijk, wat nu? En na het lezen: Wat heb ik geleerd en ben ik tevreden? Komt een leerling tot de conclusie dat het moeilijk wordt, dan zijn er voor hem voor twee situaties strategieën in de vorm van denkstappen op schrift gezet: de aanpak van moeilijke woorden en de aanpak van verwijsrelaties. Hiervoor is (in sterk vereenvoudigde vorm) gebruik gemaakt van het 'Stappenplan moeilijke woorden' (Lijmbach, Hacquebord & Galema, 1990) en het 'Programma verwijswoorden' (Aarnoutse, Roelofs & Driessen, 1989).

4.6 Resultaten en bespreking

De observatiegegevens zijn per les gecodeerd en verwerkt. In Tabel 1 wordt een overzicht gegeven van de verdeling van de lestijd in procenten. De eerste kolom geeft de gemiddelde tijdbesteding van vijf leerkrachten weer tijdens de 20 lessen voor de training (voormeting) en kolom twee geeft een overzicht van de 18 lessen na de training (nameting).

De gegevens van de voormeting geven aan dat deze leerkrachten niet veel afwijken van hetgeen Weterings en Aarnoutse (1988) en Aarnoutse (1991) reeds gevonden hebben over het lesgedrag in het basisonderwijs. Van de totale tijd tijdens lessen in begrijpend lezen wordt betrekkelijk weinig instructie gegeven. Aarnoutse (1991) vond dat leerkrachten 17% van de tijd besteden aan klassikale instructie. De meest passende vergelijking is hier wellicht een sommering van de huidige gegevens over categorieën 5 tot en met 10 en deze levert een som van 25.5% van de tijd. Dus er is enige overeenkomst tussen de huidige gegevens en eerdere resultaten van Aarnoutse (1991). Voor

Tabel 1

De gemiddelde verdeling van de lestijd (en standaarddeviatie) van 5 leerkrachten (in procenten van de totale tijd) over verschillende activiteiten tijdens 20 lessen voor en 18 na de training, en de t-waarde (df=4, eenzijdig getoetst) voor het verschil tussen voor- en nameting

Observatiecategorie	voormeting		nameting		*t*-toets	*p*-waarde
Algemeen (1-4)						
1 niet lesgericht	3.23	4.69	1.82	0.92	.80	.23
2 algemene procedures	50.33	11.21	31.36	7.79	5.02	.00
3 technisch lezen	6.96	5.53	8.53	7.35	.56	.30
4 spelling/syntaxis	1.61	2.87	0.03	0.07	1.23	.14
subtotaal	*62.13*		*41.74*			
Begrijpend lezen (5-10)						
5 instructie: voorbereiden	8.98	4.47	18.53	4.20	3.03	.02
6 instructie: voorspellen	0.18	.24	4.56	4.20	2.29	.04
7 instructie: conclusies	4.81	3.42	16.24	2.24	8.44	.00
8 instructie: tekstkenmerken	2.31	4.01	0.18	0.41	1.15	.16
9 instructie: studievaardigheden	1.34	2.83	0.61	0.62	.63	.28
10 luisteren	7.90	1.99	14.01	4.18	3.00	.02
subtotaal	*25.52*		*54.13*			
Overige (11-12)						
11 resultaatgericht begeleiden	11.16	5.16	2.23	2.10	4.30	.01
12 leerprocesgericht waarderen	1.19	1.82	1.90	0.84	1.67	.09
subtotaal	*12.35*		*4.13*			

deze vijf lom-leerkrachten moet geconcludeerd worden dat er blijkbaar geen grote verschillen zijn tussen basis- en lom-onderwijs in de tijd die er voor daadwerkelijke instructie wordt uitgetrokken.

Als deze huidige praktijk door middel van de kwalitatieve observatieschaal met een min of meer ideale situatie wordt vergeleken, dan is het oordeel weinig positief. In Tabel 2 worden de gemiddelde scores op de waarderingschaal gegeven (elke subschaal loopt van 0 naar 4). Met een maximum van 44 voor de totale schaal wordt voor de lessen tijdens de voormeting gemiddeld 6.95 gescoord. Blijkbaar worden de waargenomen lessen vanuit het oogpunt van (ideale) directe instructie in begrijpend lezen niet als kwalitatief hoogwaardig beoordeeld. Overigens komt dit cijfer redelijk overeen met de uitkomst van het onderzoek van Duffy e.a. (1987), die (voor de training) een gemiddelde van 11.45 rapporteert. Kortom, zowel de schattingen van de tijdbesteding als de kwalitatieve beoordeling tijdens de voormeting geven aanleiding tot het vermoeden dat er niet erg veel directe instructie in het (hier onderzochte) lom-onderwijs wordt gegeven. Is deze situatie met een training van leerkrachten te veranderen?

In zowel Tabel 1 als Tabel 2 staan de gegevens van de voor- en nameting naast elkaar om het effect van de training direct te kunnen nagaan. In de tabellen is door middel van de *t*-waarde (van gepaarde *t*-toetsen, eenzijdig, $df = 4$) aangegeven op welke categorieën er een significant verschil is opgetreden als gevolg van de training. Duidelijk is in Tabel 1 te zien dat de tijd die wordt besteed aan algemene procedures en aan resultaatgerichte opmerkingen aanzienlijk is afgenomen. Daarentegen is de tijd die voor instructie in begrijpend lezen wordt uitgetrokken nagenoeg verdubbeld (van 25 naar 54%). Daarbij wordt veel tijd besteed aan het trekken van conclusies en zijn de leerlingen actiever in het hardop verwoorden van leesstrategieën en hun tekstaanpak, zoals blijkt uit de toegenomen tijd waarin de leerkracht hiernaar luistert (categorie 10). Tijdens het observeren van de lessen (nameting) viel al op dat de leerkrachten veel meer vragen aan de leerlingen stellen. Dit gebeurde spontaan en lijkt inherent aan de overdracht van de leesstrategieën. Vragen als: "Waar moet je ook alweer aan denken voordat je gaat lezen? Wat doe je als je een woord niet begrijpt? Waar denk je dat dit stukje over zal gaan?" worden voortdurend gesteld. De leerlingen verwoorden hun aanpak en de leerkracht luistert. Ook wordt er meer tijd

Tabel 2

Gemiddelde waardering voor de kwaliteit van de instructie in punten per categorie (range 0 - 4) met standaarddeviaties voor respectievelijk de 20 en 18 lessen voor en na de training van 5 leerkrachten en de t-waarde (eenzijdig getoetst) voor het verschil tussen voor- en nameting

Observatiecategorie	voormeting		nameting		*t*	*p*
Informatie over de strategie						
1 Doel	0.15	0.22	1.53	0.58	6.43	.00
2 Bruikbaarheid	0.10	0.14	1.68	1.23	3.21	.02
3 Hoe te kiezen	0.00	0.00	1.38	0.44	7.04	.00
4 Hoe toe te passen	0.25	0.35	2.40	0.58	7.98	.00
Lesopbouw						
5 Introductie	1.05	0.80	2.25	0.90	2.64	.03
6 Modelling	0.10	0.14	1.80	0.43	8.03	.00
7 Verminderen van hulp	0.75	0.35	2.80	0.33	16.74	.00
8 Vragen naar oploss. proces	1.05	0.45	2.43	0.70	5.40	.00
9 Les afsluiten	1.15	0.78	1.73	0.92	3.50	.01
Samenhang						
10 in de les	1.75	1.12	3.02	0.27	3.10	.02
11 met andere lessen	0.60	0.45	1.32	0.58	2.71	.03
Totaal	*6.95*		*22.35*			

besteed aan het voorbereiden van de tekst. Dit is echter voornamelijk toe te schrijven aan het activeren van voorkennis, zowel over de tekstinhoud als ook over de verschillende leesstrategieën en dit mondt nogal eens uit in langdurige klassegesprekken. Niet voor alle categorieën is er evenveel effect zichtbaar. Instructie in het gebruik van tekstkenmerken en het toepassen van studievaardigheden is niet toegenomen (maar ook niet significant verminderd). Dit kan te maken hebben met de fasering van de lessen of met het materiaal waarmee geoefend wordt. Op de tijdstippen van observatie waren nog niet alle strategieën behandeld tijdens de lessen.

In Tabel 2 blijkt dat op de waarderingschaal voor alle categorieën een positief en statistisch significant effect van de training is gevonden. De gemiddelde waardering stijgt na de training naar een totaal van ruim 22, dat is meer dan een verdrievoudiging van de beoordeling van de beginsituatie. Deze toename is groter dan in het onderzoek van Duffy e.a. (1987) is aangetroffen; zij vonden (in de experimentele groep) een stijging van 13.40 naar 17.7 - 21.30 (er waren meerdere na-observaties) als gevolg van een training. Maar vanwege verschillen in omvang van de proefgroep, onderzoeksopzet, en de vorm en inhoud van de cursus kan niet van een echte replicatie worden gesproken.

Om de samenhang na te gaan tussen de hoeveelheid tijd die leerkrachten aan instructie besteden en de totale beoordeling van de leskwaliteit zijn in Tabel 3 per leerkracht de kwantitatieve en kwalitatieve resultaten op beide meetmomenten naast elkaar gezet. Het tijdsdeel betreft de proportie tijd die een leerkracht tijdens een les gemiddeld aan begrijpend lezen besteedt (de som van de categorieën 5 tot en met 10 uit Tabel 1) en de waardering is het puntenaantal dat een leerkracht gemiddeld behaalt op de waarderingschaal, uitgedrukt als percentage van het maximaal (is 44) te behalen puntenaantal.

Tabel 3

De tijdbesteding en de totale waardering van de mate van directe instructie in begrijpend lezen per leerkracht en moment van meting (voor of na de training, uitgedrukt in %)

	voormeting		nameting	
Leerkracht	*tijdsdeel*	*waardering*	*tijdsdeel*	*waardering*
1	22.3	19.3	53.2	55.3
2	11.4	18.2	50.3	62.9
3	19.2	10.2	54.0	36.4
4	38.8	6.8	53.5	42.6
5	35.8	24.4	59.8	56.8

Uit de derde tabel blijkt in de eerste plaats dat de positieve veranderingen in lesgedrag ten gevolge van de training voor alle leerkrachten gelden. Ten tweede kan worden opgemerkt dat deze toename voor beide observatieschalen te zien is. Klaarblijkelijk kunnen beide instrumenten heel goed het relevante lesgedrag registreren dat door een training van instructie in begrijpend lezen beïnvloed wordt. Maar ten derde blijkt bij een analyse binnen de voormeting of de nameting dat een relatief flinke tijdbesteding niet hoeft samen te gaan met een hoge kwaliteit. Een voorbeeld: tijdens de voormeting besteedt leerkracht 4 relatief de meeste tijd aan instructie in begrijpend lezen (38.8% van de lestijd), maar scoort het laagst als het gaat om een waardering van de totale les (van de 44 te behalen punten behaalt hij er gemiddeld 3, dit is 6.8% van 44). Een tweede voorbeeld zien we tijdens de nameting bij de leerkrachten 3 en 4: na de training blijken zij ongeveer hetzelfde percentage van de lestijd te besteden aan instructie als de andere leekrachten, maar de kwaliteit van de lessen blijft bij leerkrachten 3 en 4 duidelijk achter in vergelijking met de anderen. De conclusie moet zijn dat het patroon van tijdbesteding niet overeenkomt met een beoordeling van de kwaliteit van de instructie.

4.7 Discussie

De bevindingen uit het eerste onderzoek lijken aan te geven dat er net zoals in het basisonderwijs ook in deze vijf groepen in het lom-onderwijs betrekkelijk weinig directe instructie in begrijpend lezen wordt gegeven. Uit het voorgaande blijken de leerkrachten in dit onderzoek na de training duidelijk meer tijd te gaan besteden aan instructie en de lessen meer te richten op overdracht van leesstrategieën en deze explicieter te onderwijzen. Ofschoon de resultaten dus bemoedigend lijken voor de hypothese dat leerkrachten in de bovenbouwgroepen van het lom-onderwijs door training meer expliciet gaan onderwijzen in leesstrategieën, zijn een paar kanttekeningen op zijn plaats.

Het zojuist beschreven onderzoek is op te kleine schaal uitgevoerd om algemene uitspraken over lom-onderwijs mogelijk te maken. Deze observatie-studie dient eerst op grotere schaal te worden herhaald en een grotere onderzoeksgroep is gewenst. Bovendien is een controlegroep noodzakelijk om te kunnen controleren op eventuele beïnvloeding van het lesgedrag door louter de (video-)observaties in de klas. Het regelmatig observeren van leeslessen kan op zich al een verbetering in lesgedrag tot gevolg hebben, doordat leerkrachten zich gaandeweg beter voorbereiden en meer aandacht besteden aan de uitvoering van de lessen. Om te kunnen nagaan of het veranderde instructiegedrag van leerkrachten over een langere periode stand houdt en ook om hen in staat te stellen meerdere leesstrategieën aan bod te laten komen, lijkt het vervolgens raadzaam de observaties na de training enigszins te spreiden, zodat leerkrachten ook meer tijd hebben om het geleerde in praktijk te brengen. Daarnaast is het echter aannemelijk dat de fase van expliciete en intensieve instructie vooral aan het begin van een nieuw leerproces gegeven wordt, namelijk bij de introductie en behandeling van een nieuwe strategie. Zodra de leerlingen zelfstandig werken en zonder verdere hulp leesstrategieën om teksten te begrijpen gebruiken, is het leerkrachtgedrag wellicht minder directief en overwegend superviserend. Deze situatie kan enige tijd voortduren alvorens nieuwe leerstof wordt geïntroduceerd. Aangezien het lastig is vooraf te weten op welke moment in deze cyclus men observeert is ook dit een reden om observaties van lessen zoveel mogelijk te spreiden.

Ten slotte dient aandacht te worden geschonken aan de verschillende resultaten van het observatie-categorieënsysteem en de waarderingschaal. Zoals uit Tabel 3 is gebleken komen de verschillen tussen leerkrachten op deze twee schalen binnen de voormeting en binnen de nameting niet geheel overeen. Hoewel beide observatiesystemen pretenderen een middel te zijn om de kwaliteit van lessen te beschrijven rijst bij uiteenlopende resultaten de vraag aan welke de voorkeur moet worden gegeven of het meeste belang moet worden gehecht. De time-samplingprocedure levert een schatting van het patroon van de tijdbesteding en de waarderingschaal een totale beoordeling van de kwaliteit van les. De betrouwbaarheid van beide schalen is buiten kwestie (zie 4.4). Maar de validiteit is vatbaar voor kritiek. Aan de validiteit van de time-sampling procedure wordt vanwege het niet continue karakter van de observaties wel

eens getwijfeld (Mann, Ten Have, Plunkett & Meisels, 1991). De schatting van werkelijke tijdbesteding en individuele verschillen hierin kan tamelijk onnauwkeurig zijn als er veel kortdurende gebeurtenissen zijn of als er gedragscategorieën voorkomen die principieel verschillen in duur. Dat is inderdaad vaak wel het geval in een lessituatie. Het meest belangrijke is echter de vraag wat men zou willen vastleggen. Zoals eerder opgemerkt is de bestede tijd op zich onvoldoende informatief. Slechts als de veronderstelling waar is dat de hoeveelheid tijd werkelijk correspondeert met de intensiteit en kwaliteit van de instructie, kan een tijdsregistratiesysteem nuttig blijken. Zowel om inhoudelijke redenen als om de empirische discrepantie van individuele verschillen tussen leerkrachten in dit eerste onderzoek, zijn we geneigd deze veronderstelling niet aannemelijk te achten.

Maar de resultaten van dit eerste onderzoek tonen eveneens aan dat het effect van een training uitstekend door een betrouwbaar gebleken waarderingschaal kan worden vastgelegd. Omdat de procedure van tijdsregistratie zeer tijdrovend is, videogebruik ofwel meerdere observatoren in de klas vereist, en dus bij onderzoek op grotere schaal een grote materiële investering vergt, verdient het onze voorkeur om in verder onderzoek uitsluitend de waarderingschaal te gebruiken.

5 Onderzoek II

5.1 Vraagstellingen

In dit tweede onderzoek werd getracht de volgende vragen te beantwoorden: (1) Kunnen de bevindingen over het lesgedrag zoals in een kleine groep in het eerste onderzoek gevonden, gerepliceerd worden in een grotere groep leerkrachten in het lom-onderwijs, (2) kan het eerder gevonden positieve effect van een training op het lesgedrag van leerkrachten opnieuw worden gevonden in een nieuwe, grotere groep en (3) is het effect op het lesgedrag uitsluitend toe te schrijven aan een training of is regelmatige observatie (in een controlegroep) al voldoende aanleiding tot verbetering van de lespraktijk?

5.2 Onderzoeksgroep

De onderzoeksgroep bestaat uit 18 leerkrachten van 10 verschillende scholen, 10 leerkrachten ontvangen een training, 8 leerkrachten vormen de controlegroep. De voornaamste eis voor (vrijwillige) deelname was dat ze les gaven aan eindgroepen lom-onderwijs en voldoende tijd voor het onderzoek beschikbaar konden stellen. Vanwege de betrekkelijk kleine groep gegadigden (binnen een straal van ca. 80 km. rondom Amsterdam) is afgezien van verdere matching op leerkrachtkenmerken. De indeling in experimentele en controlegroep is vooral gebaseerd op eigen voorkeur van de deelnemers, omdat onvoldoende leerkrachten bereid bleken deel te nemen zonder vooraf te weten of ze bij de jaarplanning van hun activiteiten wel of geen rekening moesten houden met ingrijpende inhoudelijke veranderingen in de lessen begrijpend lezen en met het vrijroosteren van woensdagmiddagen voor deelname aan de training.

De 18 deelnemers (16 mannen en 2 vrouwen) blijken allen een ruime onderwijservaring te hebben, variërend van 5 tot 25 jaar, waarvan gemiddeld 9.4 jaar ervaring in het lom-onderwijs. Slechts twee leerkrachten blijken in de laatste vijf jaar een studiedag of cursus van een schooladviesdienst te hebben gevolgd over het vak begrijpend lezen. Beide leerkrachten maken deel uit van de experimentele groep. Ook bij drie leerkrachten uit de controlegroep is naar onze mening meer kennis over begrijpend lezen aanwezig. Zij gebruiken een (nieuwe) methode waarin uitdrukkelijk aandacht wordt besteed aan het gebruik van leesstrategieën zoals de betekenis van onbekende woorden afleiden uit de context; het afleiden van de hoofdgedachte; het formuleren en toetsen van verwachtingen over de inhoud van de tekst en door eigen associaties gegevens bedenken die impliciet in de tekst besloten liggen (de methode Taalkabaal). Naast de methode Taalkabaal blijken er zeven verschillende (oudere) methoden voor begrijpend lezen te worden gehanteerd met daarnaast gebruik van taalmethoden en technische leesmethoden. Verder blijken twee leerkrachten (één uit de experimentele en één uit de controlegroep) geen enkele methode voor begrijpend lezen te gebruiken. Gemiddeld geven de leerkrachten 1 à 2 lessen begrijpend lezen per week.

5.3 Procedure

Elke leerkracht is tijdens de voormeting gedurende ruim twee maanden in het najaar vijf maal geobserveerd (één les van één leerkracht viel echter door omstandigheden uit). In januari zijn de leerkrachten uit de experimentele groep in een cursus getraind in de directe instructie van begrijpend leesstrategieën (zie 4.5). Op basis van ervaringen uit Onderzoek 1 heeft de cursus de volgende uitbreidingen ondergaan: (a) Het aantal voorbeeldlessen is verdriedubbeld, van 6 naar 19, waarbij een nieuw onderwerp is toegevoegd, namelijk lessen over tekststructuur (het onderscheid tussen zin en regel; interpuncties; alinea-indeling; titel en sub-titels). Het door de leerkrachten zélf laten schrijven van een les analoog aan het voorbeeld is afgeschaft, omdat deze opdracht tijdens de cursus wel tot veel discussie aanleiding gaf maar nauwelijks tot concrete resultaten leidde. (b) Het aantal videobeelden ter demonstratie van instructie in leesstrategieën kon worden uitgebreid door fragmenten te nemen uit de observaties tijdens de nameting van het vorige onderzoek. (c) Het activeren van voorkennis bij leerlingen werd voorts strakker voorgestructureerd met behulp van semantische schema's (Cooper, 1986). De leerlingen en dus ook de leerkrachten werd geleerd om bij een onderwerp uit de titel "W"-vragen te stellen: wie, wat, waar, wanneer, waarom etc. om snel een context te kunnen schetsen zonder al te veel tijdrovende uitweidingen. (d) Verder zijn vier van de vijf wekelijkse trainingsbijeenkomsten in één maand gegeven en is de vijfde bijeenkomst twee maanden later gehouden. Deze ruimere periode tussen de voorlaatste en laatste bijeenkomst werd gekozen om leerkrachten gelegenheid te geven ook nog ervaringen uit te wisselen als ze al wat langere tijd met de nieuwe instructievorm aan het werk zijn. In de tussenliggende periode worden wekelijks lesevaluaties schriftelijk aan de onderzoekers gerapporteerd. Ter ondersteuning en motivatie van de leerkrachten retourneren de onderzoekers samenvattingen van deze praktijkervaringen naar de leerkrachten.

Vier weken na de training is begonnen met een tweede observatieronde. Bij elke leerkracht is vier maal, met een interval van twee weken, een les in begrijpend lezen geobserveerd. Daarbij is aan de leerkrachten geen feedback gegeven over de betreffende lessen.

Als maat voor de kwaliteit van de instructie is gebruik gemaakt van de waarderingschaal uit het vorige onderzoek (zie 4.4). De lessen van leerkrachten zijn met behulp van dasspeldmicrofoons en memorecorders vastgelegd en de opnames werden aangevuld met een schriftelijk observatieverslag. Naast de verbale uitleg van de leerkracht, die nadien zoveel mogelijk letterlijk wordt geprotocolleerd, worden voor een goed begrip van het lesverloop in het verslag ook de tijd, het lesmateriaal en de achtereenvolgende activiteiten van de leerkracht beschreven. Zo werd bijvoorbeeld genoteerd hoelang een leerkracht klassikaal instructie gaf, na hoeveel tijd er een overgang plaatsvond naar zelfstandig werken, op welk moment de leerkracht een leerling hulp bood, e.d. Op grond van de audio-opname en het verslag wordt achteraf de kwaliteit van de instructie beoordeeld met de waarderingschaal.

Vijf observatoren werden intensief getraind in het protocolleren en beoordelen van de instructie-kwaliteit van lessen in begrijpend lezen. Daarbij werd gebruik gemaakt van videoopnamen uit het eerste onderzoek. Beelden van de voormeting (voor de training in strategie-instructie) en de nameting werden door elkaar gepresenteerd, zonder de observatoren hierover te informeren. Elke beoordeelde les werd nabesproken en waar nodig werden afspraken over de beoordelingswijze bijgesteld. Na een vijftal trainingsessies van twee tot drie uur werden de observatoren vijf willekeurige op video opgenomen lessen ter beoordeling voorgelegd. De betrouwbaarheid bleek voldoende om aan de observaties in de klas te kunnen beginnen (Goodman-Kruskal's Gamma .88).

Tijdens de observatieperioden zijn de observatoren niet meer aanvullend getraind. Veertig procent van de lessen tijdens de voormeting is met behulp van een roulatiesysteem door drie beoordelaars afzonderlijk en onafhankelijk van elkaar gescoord, steeds in wisselende samenstelling, zodat de betrouwbaarheid van elk beoordelingspaar bepaald kon worden. Van de lessen in de nameting is 65% door twee van de vijf beoordelaars gescoord, ook weer in wisselende samenstelling. Hoewel de beoordelaars niet werden geïnformeerd over de samenstelling van de experimentele en controlegroep en ze cassettebandjes en protocollen kregen

met slechts de naam van de leerkracht, werd hun tijdens de nameting toch regelmatig onthuld dat ze te maken hadden met een getrainde leerkracht. Want tijdens het lesgeven spraken (uiteraard slechts) de leerkrachten uit de experimentele conditie vaak over 'de nieuwe stappenkaart', of 'de nieuwe lessen' of 'de lessen van Sophie' (voornaam van de eerste auteur van dit rapport). Uit controle beoordelingen bleek deze informatie overigens de betrouwbaarheid van de beoordeling niet aan te tasten. De interbeoordelaar-betrouwbaarheid bleef na de eerste observatieronde voldoende hoog (Gamma is .86), maar daalde na de tweede observatieronde enigszins (Gamma is .72).

5.4 Resultaten

5.4.1 Geobserveerde lessen

In totaal zijn er 161 lessen geobserveerd, 89 tijdens de voormeting en 72 tijdens de nameting. Uit een analyse van de protocollen blijkt, dat de lessen gemiddeld drie kwartier duren en variëren van 25 tot 80 minuten. In verband met klasoverstijgend geformeerde leesniveau-groepen, ligt de lesduur voor een groot aantal groepen tamelijk vast. Het valt op dat de algemene opzet van de lessen, wat betreft de gerichtheid van de aandacht van de leerkracht, eenvoudig in twee groepen te verdelen is, namelijk lessen met een voornamelijk individueel en lessen met een klassikaal gerichte aandacht van de leerkracht. Bij de overwegend klassikaal gerichte aandacht werken alle leerlingen aan eenzelfde oefening en wordt de instructie aan de hele klas tegelijk gegeven; in het andere geval heeft elke leerling of hebben groepjes leerlingen eigen teksten voor zich en is ook de instructie individueel aangepast. Tijdens de eerste observatieronde ligt de verhouding tussen individuele en klassikale gerichtheid bij de experimentele groep op respectievelijk 18 en 70% en bij de controlegroep op 49 en 51%. In de resterende 12% van de geobserveerde lessen bij de experimentele groep wordt de volledige lestijd besteed aan onderwerpen die niet zo duidelijk onder 'instructie in tekstbegrip' vallen, zoals het bespreken van een wandplaat zonder tekst, het leren zoeken in een encyclopedie en het maken van een toets. Tijdens de tweede observatieronde gaat de experimentele groep voor 100% klassikaal gerichte instructie geven conform de opzet van de voorbeeldlessen en blijft de verhouding bij de controlegroep bij benadering gelijk (47% individueel gericht en 53% klassikaal).

5.4.2 Kwaliteit van de lessen voor de training

De verwachting was dat er door de 18 leerkrachten tijdens de lessen (voorafgaand aan de training) nauwelijks directe instructie in leesstrategieën wordt gegeven. De gegevens op de waarderingschaal voor de 89 lessen tijdens de observaties in de voormeting geven aan dat er, net als in het eerste onderzoek, geen hoge waarderingsscores worden gegeven. In Tabel 4 staan de gemiddelden voor de 18 leerkrachten uitgesplitst voor de experimentele en de controlegroep. Zoals uit deze tabel blijkt behalen de leerkrachten in de experimentele groep als totale score op de schaal gemiddeld nog geen kwart (7.88 punten) van het totale aantal te behalen punten op de waarderingschaal (44 punten). De leerkrachten in de controlegroep halen wel iets hogere scores (gemiddeld 10.41), maar dit verschilt niet significant van de experimentele groep ($F(1,16) = 1.21, p = .29$). Met name twee leerkrachten uit de controlegroep zijn overigens voor deze hogere score verantwoordelijk. Zij gebruiken de methode Taalkabaal en behandelen uitdrukkelijk leesstrategieën in de lessen; deze twee behalen de hoogste totaalscores tijdens de voormetingen, namelijk rond de 20 (zie ook leerkrachten 11 en 12 in Figuur 1; de derde gebruiker van dezelfde methode is leerkracht 13 in Figuur 1).

Uit Tabel 4 blijkt weliswaar dat we te maken hebben met goede en ervaren leerkrachten, want ze zijn goed in staat om samenhang in de les te brengen (categorie 10), maar de waardering voor alle overige observatie-categorieën, die vooral ook op de specifieke instructie in begrijpend lezen betrekking hebben, bereikt verder nergens de waarde 2 (dat is het midden van de schaal 0 - 4). Hoewel de scores iets hoger zijn dan in het eerdere onderzoek, kan op grond van deze gegevens worden geconcludeerd, dat een directe en expliciete instructie in begrijpend leesstrategieën niet overtuigend voorkomt in de lessen van deze groep van 18 leerkrachten tijdens de voormeting.

Uit de observatie-protocollen van de 89 lessen in de voormeting is verder op te maken dat tijdens de lessen nogal de nadruk gelegd wordt

Tabel 4

Gemiddelde waarderingscores (en standaarddeviatie) van het lesgedrag van 18 lom-leerkrachten voor (89 lessen) en na (72 lessen) de training uitgesplitst voor de experimentele en controlegroep

	experimentele groep (10 leerkrachten) voormeting		nameting				controlegroep (8 leerkrachten) voormeting		nameting			
	m	*sd*	*m*	*sd*	*t*	*p*	*m*	*sd*	*m*	*sd*	*t*	*p*
Informatie over de strategie												
1 Doel	0.22	0.22	2.47	0.59	12.34	<0.01	0.57	0.57	0.66	0.80	0.61	0.28
2 Bruikbaarheid	0.24	0.50	3.10	0.50	12.41	<0.01	0.72	0.58	0.94	0.80	1.21	0.13
3 Hoe te kiezen	0.12	0.19	2.15	0.54	10.69	<0.01	0.23	0.23	0.59	0.33	4.77	<0.01
4 Hoe toe te passen	0.36	0.44	3.28	0.38	14.32	<0.01	0.85	1.00	0.78	0.71	0.33	0.37
Lesopbouw												
5 Introductie	0.44	0.43	2.28	0.63	7.83	<0.01	0.71	0.59	0.88	0.83	1.16	0.14
6 Modelling	0.40	0.43	3.00	0.51	11.31	<0.01	1.13	1.06	1.41	1.08	0.85	0.21
7 Verminderen van hulp	0.60	0.54	2.67	0.64	6.80	<0.01	0.57	0.47	1.34	0.83	4.31	<0.01
8 Vragen oplossings-proces	1.18	0.40	2.78	0.48	9.93	<0.01	1.58	0.89	1.94	0.66	2.33	0.03
9 Les afsluiten	0.74	0.60	2.30	0.73	4.89	<0.01	0.53	0.44	0.66	0.60	0.94	0.19
Samenhang												
10 in de les	2.42	0.37	3.58	0.33	6.36	<0.01	2.18	1.22	2.91	0.13	1.84	0.05
11 met andere lessen	1.16	0.55	2.38	0.53	8.75	<0.01	1.34	0.40	1.28	0.89	0.22	0.42
Totaal	*7.88*		*29.98*				*10.41*		*13.38*			

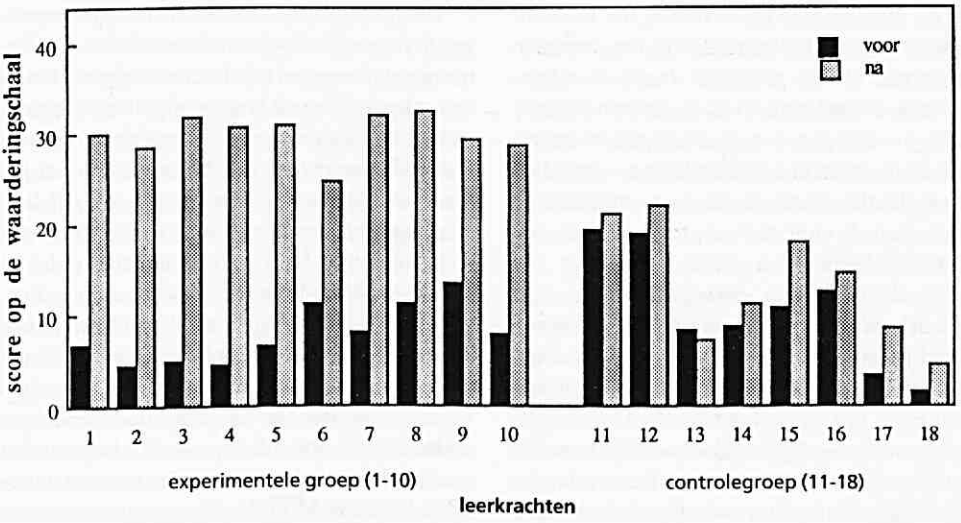
op het maken en bespreken van de opdrachten na het zelfstandig of klassikaal lezen van een tekst. De onderwerpen die aan de orde komen, zoals het samenvatten, hoofdzaken afleiden, verwijsrelaties aangeven, betekenis van onbekende woorden zoeken en voorkennis activeren, worden overwegend als een vaardigheid behandeld met de nadruk op het verwoorden van het juiste antwoord in plaats van het te volgen denkproces.

5.4.3 Effect van training

In Tabel 4 staan ook de gegevens van de observaties nadat aan de experimentele groep leerkrachten een training is gegeven. Uit de gegevens van de nameting blijkt duidelijk dat leerkrachten die getraind worden in directe instructie een hogere waardering behalen voor lessen in begrijpend lezen dan hun niet getrainde collega's. In de loop van de tijd zijn beide groepen vooruitgegaan, maar alleen de getrainde leerkrachten laten een grotere vooruitgang zien die uitsluitend geïnterpreerd kan worden als het positieve effect van de training. Een significant interactie-effect tussen observatieperiode en leerkrachtengroep, $F(1,16) = 126.48$, $p < .01$, geeft aan dat de kans dat er geen grotere vooruitgang is van de experimentele groep leerkrachten onwaarschijnlijk klein geacht mag worden. In Tabel 4 blijkt ook dat de experimentele groep tijdens de nameting op alle afzonderlijke categorieën een hogere waardering behaalt dan tijdens de voormeting, terwijl dit voor de controlegroep slechts voor een enkele categorie het geval is.

Een ander beeld van het effect van de training is te verkrijgen door de beoordeling van de lessen per leerkracht weer te geven. In Figuur 1 is voor alle 18 leerkrachten, de eerste 10 vormen de experimentele groep en de tweede 8 de controlegroep, de gemiddelde totaalbeoordeling gegeven van de lessen die in de periode voor en na de training zijn gegeven. Uit de grafiek blijkt overduidelijk dat de eerder vermelde effecten niet slechts op groepsgemiddelden zijn gebaseerd, maar dat het effect van de training voor alle leerkrachten opgeld doet.

Een interessant gegeven is ten slotte nog dat er een hoge correlatie bestaat ($r = + 0.93$) tussen de scores van de controlegroep op de voor- en nameting, maar dat de resultaten van de experimentele groep veel minder en zelfs negatief correleren ($r = - 0.24$). Uit deze correlaties kan worden opgemaakt, dat aanvankelijke verschillen tussen de experimentele leerkrachten niet of nauwelijk iets te maken hebben met de verschillen na de training, zelfs zo dat deze verschillen in een andere richting komen te liggen. Bij de controlegroep ligt dit duidelijk anders. Hier blijven de onderlinge verschillen tussen leerkrachten grotendeels hetzelfde. Hoewel gemiddeld de waardering van het instructieve karakter van de lessen iets is toegenomen (van 10.41 naar 13.38), blijven de aanvankelijke verschillen tussen deze leerkrachten nagenoeg onaangeroerd. We zouden daaruit willen con-



Figuur 1. Gemiddelde scores op de waarderingschaal per leerkracht in de periode voor en na de training (leerkracht 1-10 hebben wel en 11-18 hebben geen training ontvangen)

cluderen dat er een bepaalde mate van stabiliteit in het lesgedrag gedurende het schooljaar aanwezig lijkt te zijn. De hoge correlatie in de controlegroep is overigens tevens een pleit voor de betrouwbaarheid van de observaties: had men tijdens de nameting bijvoorbeeld op willekeurige wijze strenger of milder geoordeeld, dan waren er waarschijnlijk al snel (kleine) verschillen in rangorde te zien geweest bij de controlegroep.

6 Conclusies en discussie

De eerste vraag in dit onderzoek was of het instructiegedrag van leerkrachten in het lom-onderwijs tijdens lessen in begrijpend lezen vergelijkbaar is met dat van collega's in het basisonderwijs. Deze vraag kan in grote lijnen bevestigend beantwoord worden voor de 23 leerkrachten die aan een van beide onderzoeken hebben deelgenomen. Wat betreft de vraag naar tijdbesteding kunnen we uit het eerste onderzoek concluderen, dat de helft van de lestijd besteed wordt aan algemene procedures (management, supervisie en gedragsregulatie) en ongeveer eenvijfde van de tijd aan beoordelingen (resultaatgerichte begeleiding) en het luisteren naar antwoorden van leerlingen op vragen en opdrachten. Aan instructie wordt weliswaar enige lestijd besteed, maar bij de geobserveerde groepen wordt nauwelijks tot geen informatie over leesstrategieën expliciet onderwezen. Dit blijkt uit de kwalitatieve beoordeling van de lessen in beide onderzoeken.

Wat betreft de vorm van de lessen kan opgemerkt worden, dat in ca. 60% van de geobserveerde lessen bij de 18 eindgroepen van het lom-onderwijs de aandacht van de leerkracht vooral klassikaal gericht is en alle leerlingen met eenzelfde taak bezig zijn. Ook bestaat er een heldere samenhang in de les, maar de relatie met voorgaande of volgende lessen wordt zelden duidelijk voor de leerlingen verwoord. Het directe instructiemodel is in de lessen niet of nauwelijks terug te vinden: introductie en modelleren van hetgeen geleerd gaat worden, het verminderen van hulp en het afsluiten van de les worden sporadisch waargenomen. In geringe mate wordt er wel gevraagd naar het gevolgde oplossingsproces bij gemaakte opdrachten.

De conclusie is dat in het lom-onderwijs doorgaans betrekkelijk weinig echte instructie in begrijpend lezen, dat wil zeggen, een expliciete en directe instructie in het gebruik van leesstrategieën, wordt gegeven. Deze slotsom moet overigens met enige voorzichtigheid worden getrokken, want gegevens over de representativiteit van de huidige onderzoeksgroep zijn niet beschikbaar. Een dergelijke te-

rughoudendheid geldt echter ook ten aanzien van de gerefereerde resultaten in het basisonderwijs, want ook daar is slechts een beperkte steekproef uit de populatie basisschoolleerkrachten bestudeerd. Gezien de consistentie van de huidige gegevens met het eerdere onderzoek en de onderling slechts kleine verschillen tussen leerkrachten in de hier onderzochte groep, kan de slotconclusie toch niet erg onwaarschijnlijk worden geacht.

Uit de twee hier gerapporteerde onderzoeken komt verder naar voren, dat training een duidelijk positief effect heeft op het lesgedrag. Leerkrachten blijken na een intensieve training gedurende vijf bijeenkomsten hun instructiegedrag in de dagelijkse klassepraktijk tamelijk ingrijpend te hebben veranderd. Deze veranderingen behelzen o.a.: a) voordat de leerlingen gaan lezen wordt expliciet aandacht besteed aan het activeren van voorkennis, het bepalen van een leesdoel en het formuleren van verwachtingen; b) tijdens en na het lezen worden verwachtingen getoetst en bijgesteld, tekst(delen) samengevat, referentiële relaties gelegd en de betekenis van onbekende woorden achterhaald. De cognitieve en metacognitieve vaardigheden en strategieën van een lezer, die een tekst grondig tracht te verstaan, worden op basis van op schrift gestelde voorbeeldlessen door de getrainde leerkrachten voorgedaan, gedemonstreerd en expliciet overgedragen. Deze veranderingen in het lesgedrag zijn in beide onderzoeken duidelijk zichtbaar, maar er moet worden opgemerkt dat in beide onderzoeken de effecten op betrekkelijk korte termijn gemeten zijn. De observaties vonden plaats op momenten dat de aan te leren vaardigheden niet alleen voor de leerlingen, maar vooral ook voor de leerkrachten nog relatief nieuw waren. Om veranderingen in lesgedrag werkelijk op waarde te kunnen schatten zou het aanbeveling verdienen om door vervolgonderzoek na te gaan hoe de instructie er in de praktijk op langere termijn uit zou zien. Overigens is het moment van observatie zoals reeds eerder opgemerkt (zie 4.7) altijd onderhevig aan mogelijke kritiek dat observaties beïnvloed kunnen worden door faseringen in de leerstof. Het geven van een bepaalde les(inhoud) kan met andere woorden meerdere lesuren bestrijken. Het enige wapen hiertegen is om meerdere keren (bij voorkeur zo frequent mogelijk) en verspreid in de tijd te observeren.

Het huidige onderzoek laat zien dat het mogelijk is om leerkrachten door middel van een training te bewegen tot effectief onderwijzen in een schoolvak waar hogere cognitieve processen en strategieën een belangrijke rol spelen. Uit eerder onderzoek is gebleken dat op een gebied waar de leerstof zich vrij gemakkelijk laat structureren en sequentiëren, bijscholing van leerkrachten in het directe instructiemodel in het algemeen wel succesvol is, maar dat het resultaat nogal wisselend is als de stof betrekking heeft op hogere strategische processen (Veenman, 1992). Het huidige onderzoek toont bovendien aan, dat een experimenteel trainingsonderzoek voor de praktijk belangrijke resultaten kan opleveren en juist door de dagelijkse klassepraktijk als uitgangspunt te nemen een hoge ecologisch validiteit heeft. Deze naturalistische oriëntatie brengt uiteraard tegelijkertijd een aantal beperkingen met zich mee voor wat betreft het onderzoeksdesign. De samenstelling en representativiteit van de onderzoeksgroep is al eerder genoemd. Ook is onvermijdelijk dat leerkrachten 'hun beste beentje zullen voorzetten' omdat er geobserveerd wordt. Toch hoeft dit het onderzoek niet nadelig te beïnvloeden. Uit Amerikaans onderzoek (Wendler, Samuels & Moore, 1989) is gebleken, dat leerkrachten onder invloed van observatoren in de klas meer tijd gaan besteden aan het individueel helpen van leerlingen tijdens het zelfstandig maken van opdrachten bij een tekst, maar dat het instructiegedrag zelf inhoudelijk niet verandert. Dit is ook heel aannemelijk als er van uitgegaan wordt, dat vaardigheden die niet worden beheerst ook niet kunnen worden getoond. De in ons onderzoek gegeven training heeft wel een aantoonbaar effect op inhoudelijke aspecten van het lesgedrag.

Een training van leerkrachten kan een belangrijke bouwsteen zijn voor het scheppen van optimale onderwijscondities voor oudere zwakke lezers in het lom-onderwijs. De mogelijke gevolgen van een training voor het actuele lesgedrag worden in de huidige onderzoeksgegevens duidelijk geïllustreerd. Een vraag is of het gewenste lesgedrag ook op langere termijn na de training uitgevoerd wordt. De laatste observaties in het huidige onderzoek waren ongeveer drie maanden na de laatste trainingsdag. Hoe de lessen na een langere periode worden

gegeven is thans nog onbekend. Op zijn minst drie zaken spelen hier een rol.

Ten eerste is aannemelijk dat veranderingen in didactisch handelen, die het gevolg zijn van ingrijpende wijzigingen in de *opvattingen* over instructie van begrijpend lezen en betrekking hebben op cognitief-strategische procedures, tamelijk permanent zijn en een grotere duurzaamheid zullen vertonen dan meer oppervlakkige wijzigingen.

Ten tweede, een belangrijke bijdrage tot de instandhouding van het gewenste didactische handelen zal vermoedelijk worden gevormd door de reacties van de leerlingen. Als leerlingen responsief zijn op de wijzigingen in lesgedrag en gaandeweg ook strategischer te werk gaan, dan zal een leerkracht vanwege dit succes eerder en langer geneigd zijn de nieuwe aanpak te blijven volgen. Op dit moment zijn we bezig na te gaan of leerlingen daadwerkelijk profiteren van de veranderingen in het lesgedrag. Gezien de gunstige resultaten die o.a. Palincsar en Brown (1989) rapporteren, mogen hier positieve effecten verwacht worden.

Ten derde, hoewel de nieuwe benadering cognitief-strategisch geaard is en in principe ook toepasbaar is onafhankelijk van het lesmateriaal of de methode, hebben we gemerkt dat leerkrachten behoefte hebben aan geschikt gemaakte materialen en lessuggesties. Een probleem voor het lom-onderwijs is namelijk dat er nog relatief weinig lesmateriaal is dat voor zwakke lezers is aangepast, dat wil zeggen, teksten die technisch gezien niet al te moeilijk zijn, maar toch geschikt zijn om begrijpend lezen mee te oefenen. Een passende reeks materialen en uitgewerkte lessuggesties kan leerkrachten voor langere tijd ondersteuning bieden en ze blijvend herinneren aan en oproepen tot de toepassing van de geleerde aanpak. Dat een methode invloed heeft, is te zien aan de twee leerkrachten in de controlegroep die reeds een op de markt gebrachte methode gebruiken waarin veel recente inzichten over een nieuwe stijl van instructie in begrijpend lezen zijn verwerkt. Er komen de laatste jaren meer nieuwe methoden op de markt en het zou interessant zijn om door middel van observatiestudies na te gaan of het lesgedrag werkelijk verandert indien deze methoden worden gebruikt.

Literatuur

Aarnoutse, C. A. J. (1991). Begrijpend lezen in het basisonderwijs. In P. Reitsma & A. M. A. Walraven (red.), *Instructie in begrijpend lezen* (pp. 129-142). Delft: Eburon.

Aarnoutse, C. A. J., Roelofs, E. C., & Driessen, J. W. (1989). *Programma verwijswoorden*. Nijmegen: Berkhout.

Barr, R., & Sadow, M. W. (1989). Influence of basal programs on fourth-grade reading instruction. *Reading Research Quarterly, 24*, 44-71.

Boonman, J. H., Kok, W. A. M., & Haan, M. de (1988). *Video-programma "Lezen is denken"*. Utrecht/ Hoevelaken: ISOR/CPS.

Cooper, J. D. (1986). *Improving reading comprehension*. Boston: Houghton Mifflin Company.

Creemers, B. P. M. (1991). *Effectieve instructie. Een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's Gravenhage: SVO.

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M. S., Vavrus, L. G., Wesselman, R., Putnam, J., & Bassiri, D. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly, 22*, 347-368.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*, 481-533.

Durkin, D. (1981). Reading comprehension instruction in five basal reader series. *Reading Research Quarterly, 16*(4), 515-544.

Glopper, K. de, & Otter, M. E. (1992). Leesprestaties Nederlandse kinderen vallen tegen. *Didaktief, 22* (10), 4-6.

Lijmbach, B. V. J. M., Hacquebord, H. I., & Galema, K. (1990). *Weet wat je leest*. Groningen: Dijkstra.

Mann, J., Have, T. ten, Plunkett, J. W., & Meisels, S. J. (1991). Time sampling: A methodological critique. *Child Development, 62*, 227-241.

Otter, M. E., Oostdam, R. J., & Vergeer, M. M. (1988). *Voorstudie Periodieke Peiling van het Onderwijsniveau in het Speciaal Onderwijs. Deel IV: Het onderwijsaanbod voor taal op lom-scholen*. Amsterdam: SCO.

Palincsar, A. S., & Brown, A. L. (1989). Instruction for selfregulated reading. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 19-39), Yearbook of the Association for Supervision and Curriculum Development. Alexandria: ASCR.

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*, 1239-1252.

Paris, S. G., & Oka, E. R. (1986). Children's reading strategies, metacognition and motivation. *Developmental Review, 6*, 25-56.

Paris, S. G., Wasik, B. A., & Turner, J. C. (1991). The development of strategic readers. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of research on reading, vol. II* (pp. 609-640). New York: Longman.

Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of research on reading, vol. II* (pp. 815-860). New York: Longman.

Pearson, P. D., Dole, J., Duffy, G., & Roehler, L. (in press). Developing expertise in reading comprehension: What should be taught? How should it be taught? In J. Farstrup & S. J. Samuels (Eds.), *What research has to say to the teacher of reading* (2nd ed.). Newark, NJ: International Reading Association.

Pennings, A., Jong, T. de, & Laudy, M. (1989). Het bevorderen van het gebruik van leesstrategieën via een dialogische interventieprocedure bij kinderen met leerproblemen. *Tijdschrift voor Orthopedagogiek, 28*, 541-551.

Roehler, L. R., & Duffy, G. G. (1991). Teachers' instructional actions. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of research on reading, vol. II* (pp. 861-883). New York: Longman.

Reitsma, P., & Walraven, A. M. A. (red.) (1991). *Instructie in begrijpend lezen*. Delft: Eburon.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Sijtstra, J. (1993). *Balans van het taalonderwijs halverwege de basisschool*. Arnhem: CITO

Sliepen, S. E. (1991). Tekstbegrip bij zwakke lezers, I & II, *Jeugd in School en Wereld, 76* (2), 24-27 & *76* (3), 18-20.

Snippe, J. (1991). *In-service training voor leerkrachten*. Groningen: RION.

Tillema, H. H., & Jong, R. de (1991). Training en coaching van bekwaamheden bij leraren. In A. Pennings e.a. (red.) *Bijdragen aan pedagogisch onderzoek 1990*. Amersfoort/Leuven: ACCO.

Veenman, S. (1992). Effectieve instructie volgens het direct instructiemodel. *Pedagogische Studiën, 69*, 242-269.

Vergeer, M. M., Oostdam, R. J. & Otter, M. E. (1988). *Taalonderwijs op LOM- en MLK-scholen. Voorstudie Periodieke Peiling van het onderwijsniveau*. 's Gravenhage: SVO.

Walraven. A. M. A. (1990). *Videoband "Instructie in leesstrategieën in kleine groepen."* Amsterdam / Duivendrecht: Paedologisch Instituut.

Wendler, D., Samuels, S. J., & Moore, V. K. (1989). Comprehension instruction of award-winning teachers, teachers with master's degrees and other teachers. *Reading Research Quarterly, 24* (4), 382-401.

Weterings, A. C. E. M., & Aarnoutse, C. A. J. (1986). De praktijk van het onderwijs in begrijpend lezen. *Pedagogische Studiën, 63*, 387-400.

Manuscript aanvaard 7-6-1993

Auteurs

S. E. Sliepen is als projectmedewerkster aangesteld bij NWO en gedetacheerd op het Paedologisch Instituut voor de uitvoering van een onderzoek naar instructie in tekstbegrip op lom-scholen.
P. Reitsma is projectleider en verbonden aan het Paedologisch Instituut Amsterdam/Duivendrecht.

Correspondentie-adres: Paedologisch Instituut Amsterdam, Postbus 303, 1115 ZG Duivendrecht.

Abstract

Training teachers from special elementary schools in using direct instruction of reading comprehension strategies.

S. E. Sliepen & P. Reitsma. Pedagogische Studiën, 1993, 70, 420-444.

Recent research has shown that in normal elementary schools almost no direct instruction is provided in using strategies for reading comprehension. No results are yet available about teachers in schools for learning disabled children. Therefore, the present two studies were designed to study teacher's behaviour in actual classrooms during lessons for read-

ing comprehension. Lessons were video- or audiotaped and teachers behaviours were categorized in terms of time devoted to various activities and quality on several dimensions of instructional behaviour. The results show that almost no direct instruction in reading comprehension strategies were given.

In two subsequent studies 5 and 10 teachers respectively were trained to directly teach the flexible use of various comprehension strategies in reading. Actual reading lessons of these teachers and of a control group of eight teachers who did not receive the training, were observed both before and after the training. The results definitively show that as a result of training the teachers actually devoted much more time to directly and explicitly teaching reading comprehension strategies and improved the quality of instruction according to a direct instruction model for teaching comprehension strategies.

Appendix

Observatiecategorieën voor waardering van instructie (naar Duffy e.a., 1987)

Informatie over de strategie

1 Beoordeel hoe expliciet de leerkracht aan het begin of tijdens de les aan de leerlingen vertelt dat er een strategie geleerd gaat worden om een probleem tijdens het lezen op te lossen (*het doel van de les*).

0 De leerkracht vertelt absoluut niet wat er geleerd gaat worden. Er kan wel een opdracht gegeven worden, maar deze heeft geen betrekking op een leesstrategie ("Jullie krijgen zo dadelijk een tekst met vragen die je gaat maken").

1 Het doel van de les (de te leren strategie) wordt genoemd, maar blijft beperkt tot informatie over *wat* er geleerd gaat worden. Er wordt geen aanvullende informatie gegeven ("We gaan de hoofdgedachte leren", "We leren vandaag weer een stapje van de strategie erbij").

2 Het doel van de les (de te leren strategie) wordt beschreven in termen van *wat en hoe* er geleerd gaat worden ("We gaan leren hoe we ....").

3 Het doel van de les (de te leren strategie) wordt omschreven als een toepasbare, flexibele strategie, maar is nog niet volledig; er wordt gesproken over *wat, hoe, wanneer òf waarom* ("We gaan leren hoe... en wanneer/waarom we..").

4 De taak wordt op perfecte wijze gepresenteerd als een toepasbare flexibele strategie om tijdens het lezen een probleem op te lossen (*wat, hoe, wanneer en waarom*).

2 Beoordeel hoe expliciet de leerkracht de leerlingen *informeert over de bruikbaarheid* van de strategie tijdens het lezen.

0 De leerkracht vertelt absoluut niet waarvoor de strategie gebruikt kan worden.

1 De leerkracht noemt slechts dat de strategie in het algemeen of tijdens het lezen bruikbaar is, maar specificeert niet waarom en wanneer.

2 De bruikbaarheid wordt gerelateerd aan de toekomst ("Als je in groep 8 komt..") of er wordt in vage en algemene termen verwezen naar het waarom of wanneer in relatie tot een bepaalde tekst ("Het helpt je om informatie te krijgen..").

3 De onmiddellijke bruikbaarheid van de strategie wordt geïllustreerd door naar een voorbeeld te verwijzen, maar de leerkracht demonstreert de bruikbaarheid niet zelf aan de hand van een voorbeeld ("Als je het woord veeteelt niet kent, kan deze strategie je daarbij helpen").

4 De leerkracht demonstreert de bruikbaarheid van de strategie door deze - al hardop denkend - zelf toe te passen aan de hand van enkele concrete voorbeelden tijdens het lezen van een tekst.

3 Beoordeel hoe expliciet de leerkracht aan de leerlingen uitlegt *hoe* ze uit meerdere strategieën de juiste strategie kunnen *kiezen* als ze een probleem tijdens het lezen tegenkomen.

0 Er wordt absoluut niet gesproken over het

kiezen van een strategie om het probleem op te lossen.

1 De leerkracht noemt slechts deze strategie als mogelijkheid, of noemt naast deze strategie nog andere oplossingsmogelijkheden,maar geeft geen aanvullende informatie over de keuze.

2 De leerkracht zegt dat deze strategie gebruikt kan worden om het probleem op te lossen en geeft enige informatie over de manier waarop de meest geschikte strategie gekozen kan worden (enkele 'voors en/of tegens' worden genoemd).

3 De probleemsituatie wordt nadrukkelijk gespecificeerd en de manier waarop een geschikte strategie gekozen kan worden wordt benadrukt, maar wordt niet door de leerkracht aan de hand van voorbeelden gedemonstreerd.

4 De leerkracht demonstreert expliciet hoe je een probleem kunt herkennen en hoe een passende strategie gekozen kan worden.

4 Beoordeel hoe expliciet de leerkracht zijn leerlingen vertelt *hoe* de strategie moet worden *toegepast* tijdens het lezen van een tekst.

0 Er wordt absoluut niet uitgelegd hoe de strategie moet worden toegepast.

1 Er wordt uitleg gegeven in de vorm van regels of een procedure die uit het hoofd geleerd moet worden. Er wordt geen voorbeeld gegeven.

2 De leerkracht praat over de regel en/of procedure als een routinematige handeling, die zonder variatie toegepast moet worden en geeft voorbeelden.

3 De leerkracht vertelt dat de leerlingen de mentale stappen op een flexibele manier kunnen toepassen, maar verwijst daarbij naar voorbeelden en demonstreert dit niet zelf.

4 De leerkracht demonstreert en beschrijft hoe de leerlingen tijdens het uitvoeren van de strategie de mentale stappen op een flexibele manier kunnen toepassen.

De lesopbouw

5 Beoordeel hoe expliciet de leerkracht de les introduceert (*introductie*).

0 De leerkracht maakt geen introducerende opmerking en geeft geen overzicht van de les. Er kan wel een opdracht gegeven worden maar daarin wordt niet verteld wat er geleerd gaat worden ("Jullie krijgen zo dadelijk een tekst met vragen die je gaat maken.").

1 De leerkracht geeft aan wat er geleerd gaat worden, maar vertelt niet waarom en hoe.

2 De leerkracht geeft aan wat er geleerd moet worden, vertelt waarom *òf* hoe (maar niet allebei).

3 De leerkracht geeft aan wat, waarom en hoe geleerd gaat worden, maar geeft geen voorbeeld.

4 De leerkracht geeft aan de hand van voorbeelden informatie over de te leren vaardigheid (strategie), de tekst waarbij deze zal worden toegepast en waaraan gedacht moet worden tijdens het gebruik ervan.

6 Beoordeel hoe expliciet de leerkracht de mentale stappen aan de leerlingen voordoet tijdens het identificeren van het probleem, het selecteren van de strategie en het toepassen ervan (*modelling*).

0 Op geen enkel moment in de les doet de leerkracht voor hoe de taak moet worden aangepakt.

1 De leerkracht noemt de procedurele stappen.

2 De leerkracht demonstreert de achtereenvolgende stappen, maar maakt het denkproces voor de leerlingen niet duidelijk.

3 De leerkracht past de strategie toe en demonstreert daarbij de mentale stappen (denkt hardop), maar doet dit fragmentarisch of onvolledig of met gekunstelde voorbeelden.

4 De leerkracht demonstreert aan de hand van de teksten waarmee de leerlingen op dat moment bezig zijn, hoe de mentale stappen tijdens het uitvoeren van de strategie verlopen.

7 Beoordeel hoe expliciet de leerkracht tijdens zijn instructie van regulatie van de strategie door de leerkracht zelf overstapt naar zelfregulatie door de leerling (*verminderen van hulp*).

0 De leerkracht geeft de leerling absoluut geen gelegenheid om onder zijn leiding te oefenen.

1 De leerkracht last één oefenfase in waarin het gebruik van de strategie noodzakelijk is, maar blijft de antwoorden benadrukken in plaats van het mentale proces (een vraag-antwoord spel of een beoordelende begeleiding).

2 De leerkracht gaat in twee of meer oefenfasen van leerkrachtregulatie over naar leerlingregulatie, maar blijft de antwoorden benadrukken in plaats van het mentale proces.

3 De leerkracht gaat in twee of meer oefenfasen van leerkrachtregulatie over naar zelfcontrole door de leerling en benadrukt de mentale processen meer dan de resultaten, maar het is nog niet perfect.

4 De leerkracht creëert een aantal weloverwogen oefensituaties, die gekenmerkt worden door hulp in een vroeg stadium van de les en begeleiding van de leerling bij het toepassen van bewuste controleprocessen, door het stellen van vragen die hiernaar refereren.

8 Beoordeel in welke mate de leerkracht antwoorden van de leerlingen benadrukt waarin vertelt moet worden hoe ze aan dit antwoord gekomen zijn (*vragen naar het oplossingsproces*).

0 De leerkracht gaat niet in op of vraagt niet naar de antwoorden van de leerlingen.

1 De leerkracht benadrukt goede antwoorden, maar verlangt van de leerlingen niet dat ze vertellen hoe ze aan het antwoord gekomen zijn.

2 De leerkracht verlangt van de leerling dat deze vertelt hoe hij aan het antwoord is gekomen, maar richt de nadruk daarbij op de procedurele stappen in plaats van de wijze waarop het antwoord verkregen is.

3 De leerkracht laat leerlingen vertellen hoe ze aan het antwoord gekomen zijn (de procedurele én mentale stappen), maar vraagt daarbij niet om een toepassing en/of helpt de leerling onvoldoende om foutieve antwoorden en denkstappen te corrigeren.

4 De leerkracht vraagt voorbeeldgewijs om toelichting en verlangt van de leerling dat deze alle mentale stappen noemt tijdens het strategisch toepassen van een vaardigheid.

9 Beoordeel hoe goed de leerkracht de geobserveerde les (of het lesdeel) *afsluit*.

0 Er is absoluut geen sprake van een afronding van een lesdeel of van de les ("Jullie hebben het nu gezien, ga maar beginnen", "Het is tijd, leg je werk maar op de hoek van de bank").

1 De leerkracht rondt het lesdeel of de les af (door bijvoorbeeld de antwoorden op de gemaakte vragen te bespreken), maar vat de geleerde strategie niet samen.

2 De leerkracht geeft een gedeeltelijke samenvatting (wat en/of waarom en/of hoe).

3 De leerkracht vat samen wat er geleerd is, waarom en hoe het moet worden toegepast, maar doet dit zonder de leerlingen er in te betrekken.

4 De leerkracht geeft een uitstekend voorbeeld door samen met de leerlingen samen te vatten en/ of te evalueren, en/of door de geleerde vaardigheid in andere teksten toe te passen en/of door de leerlingen eraan te herinneren dat de strategie ook in andere situaties kan worden gebruikt.

Samenhang tussen lessen en lesonderdelen

10 Beoordeel hoe goed de leerkracht er in slaagt om samenhang *in de les* te brengen.

0 Er is absoluut geen herkenbare samenhang of opeenvolging in de les.

1 Er lijkt een logische sequentie in de les te zitten, maar hierin treden regelmatig inconsistenties en onderbrekingen op.

2 De les weerspiegelt een logische opeenvolging, maar bevat inconsistenties ten aanzien van het doel, of de opeenvolging van activiteiten verloopt niet soepel.

3 De les bevat structuur, is consistent, doelgericht en verloopt vlot, maar het is nog niet perfect. Bijvoorbeeld: de leerkracht besteedt (bewust en gepland) meer aandacht aan het samenwerkingsproces dan aan instructie in leesstrategieën; verwoordt het denkproces onvoldoende of wijkt teveel af van het algemene leesdoel (bijvoorbeeld, een les in wereldoriëntatie).

4 De leerkracht geeft een les in begrijpend lezen die een uitstekend voorbeeld is in termen van interne struktuur, consistentie, doel en soepel verloop.

11 Beoordeel hoe goed de leerkracht de samenhang met *voorgaande en volgende lessen* weergeeft.

0 De relatie met voorgaande en komende lessen is absoluut niet herkenbaar.

1 De leerkracht verwijst naar voorgaande óf volgende lessen, maar de samenhang is niet duidelijk.

2 De leerkracht verwijst naar voorgaande óf naar volgende lessen, de samenhang is duidelijk.

3 De leerkracht verwijst naar voorgaande én volgende lessen en er is sprake van samenhang, maar het is nog niet perfect.

4 De leerkracht geeft een voorbeeldles in termen van samenhang met andere lessen.

Determinanten van seksespecifieke vakkenkeuzen in HAVO en VWO

H. Dekkers*

Samenvatting

Ter beantwoording van de vraag naar de oorzaken van sekseverschillen in (exacte) vakkenkeuzen in het Voortgezet Onderwijs is een longitudinale studie uitgevoerd waarin leerlingen vanaf de brugklas tot aan hun vakkenpakketkeuze zijn gevolgd. Ieder jaar zijn de (ontwikkelingen in) attitudes, attributies, prestaties en voorgenomen vakkenkeuzen gemeten. Getracht is onder andere in een model de (interacterende) invloed van predictoren op vakkenkeuze (m.n. wiskunde en natuurkunde, en als vergelijkingsvak Frans) aan te tonen, rekening houdend met sekse- en schooltypeverschillen. Wijzigingen in bepaalde factoren (m.n. het vak leuk, interessant en gemakkelijk vinden, en geschiktheid voor het vak) hangen in dezelfde mate voor jongens als voor meisjes samen met wijzigingen in (voorgenomen) vakkenkeuzen. Andere factoren (toekomstplannen, attributie, een vak nuttig vinden voor de toekomst) vertonen wel sekseverschillen, maar beïnvloeden de vakkenkeuze niet.

Inleiding

Het emancipatiebeleid van de Nederlandse overheid is al een ruim aantal jaren gericht op het bevorderen van de economische zelfstandigheid van vrouwen; het onderwijs-emancipatiebeleid in het verlengde daarvan voornamelijk op de beroepskwalificatie-functie van het onderwijs. Ondanks kanttekeningen van onder andere de Emancipatieraad bij de eenzijdige benadering van meisjes bestaat het concrete beleid uit het bevorderen van de deelname aan exacte en technische vakken door meisjes, zowel in het beroepsonderwijs als in het voortgezet onderwijs en de vervolgopleidingen. De deelnamecijfers geven hier zeker aanleiding toe, al zou men kunnen tegenwerpen dat jongens minder vertegenwoordigd zijn bij de taalvakken, dat het tekort aan arbeidskrachten in de technische sector misschien tijdelijk is en dat 'Europa '92' juist in taalvakken geschoolde functionarissen behoeft. Voorlopig echter concentreert de overheid zich zowel in haar beleid als, daarmee samenhangend, in onderzoeksvraagstellingen op (de oorzaken van) de geringe deelname van meisjes aan exacte vakken en studierichtingen.

Dat er sekseverschillen in de keuze voor exacte vakken bestaan is ruim gedocumenteerd. Tabel 1 toont de percentages meisjes en jongens die exacte vakken kiezen in hun eindexamenpakket in het Voortgezet Onderwijs in Nederland in de laatste tien jaar.

De cijfers tonen behoorlijke verschillen tussen jongens en meisjes, alhoewel die vooral voor wiskunde de laatste jaren afnemen. Dit is in overeenstemming met de resultaten van een meta-analyse van Hyde, Fennema en Lamon (1990), die een afname van de verschillen constateren over de laatste 20 jaar. Ook Dossey, Mullis, Lindquist en Chambers (1988) tonen een afname van sekseverschillen in wiskunde in het laatste decennium; Mullis en Jenkins (1988) voegen daaraan toe dat de verschillen voor natuurkunde even groot blijven.

Vanuit het uitgangspunt dat de sekseverschillen in vakkenkeuzen in het Voortgezet Onderwijs resultaat zijn van een ontwikkelingsproces, dat zich voor een groot deel afspeelt na het basisonderwijs, dus in de eerste jaren van het Voortgezet Onderwijs, is begin 1990 een longitudinaal onderzoek gestart naar (de ontwikkeling van) factoren die een rol spelen bij de uiteindelijke vakkenpakketkeuze in het AVO/VWO. Jaarlijks is een groot aantal

* Met dank aan J. Doesborgh voor zijn ondersteuning bij de analyses

Tabel 1
Deelnemers naar gekozen exacte vakken (%)

VWO	1983		1985		1987		1989		1991	
	M	V	M	V	M	V	M	V	M	V
Wiskunde A(I)	82	57	82	59	62	54	61	58	61	61
Wiskunde B(II)	25	5	27	6	63	30	62	31	63	32
Natuurkunde	64	28	64	28	63	28	62	28	62	29
Scheikunde	57	35	56	35	52	32	48	29	46	27

HAVO	1983		1985		1987		1989		1991	
	M	V	M	V	M	V	M	V	M	V
Wiskunde	70	39	73	42	76	46	80	51	77	57
Natuurkunde	49	14	51	14	50	13	51	12	51	17
Scheikunde	47	34	46	32	42	28	40	24	37	24

Afgeleid uit: CBS-statistieken voor VWO/HAVO/MAVO

leerlingen gevolgd in hun leerprestaties en voor het onderzoek relevante attitudes. In dit artikel bespreken we de theoretische invalshoeken en opzet van dit onderzoek, en de resultaten tot en met het derde jaar van de metingen. We beperken ons tot het HAVO, waar de definitieve vakkenpakketkeuze aan het eind van het derde jaar plaatsvindt, en het VWO, waar dan nog sprake is van voorgenomen keuzen.

1 Theoretisch kader

Vanuit diverse invalshoeken is de afgelopen jaren onderzoek gedaan naar de oorzaken van seksespecifieke keuzen in het AVO/VWO. Zowel vanuit sociaal-psychologische hoek als vanuit onderwijskundige en sociologische hoek zijn verklaringen aangedragen voor de geringe belangstelling van meisjes voor vakken als wiskunde en natuurkunde. In de opzet van het onderhavige onderzoek is getracht de inzichten uit dergelijk onderzoek te *integreren*. Bovendien is gepoogd een belangrijk manco te ondervangen, namelijk het ontbreken van dataverzamelingen die recht doen aan het inzicht dat er sprake is van een *ontwikkelingsproces* dat leidt tot bepaalde keuzen.

De 'rivaliserende' verklaringen hebben enerzijds betrekking op leerlingkenmerken, anderzijds op leerkracht- en schoolkenmerken.

Variabelen op leerlingniveau vormen het hoofdbestanddeel van het onderzoek. De meeste van deze variabelen passen in een verklaringsmodel uit de sociale psychologie (waarde/verwachtingsmodel), dat Eccles (o.a. 1986) toepaste om factoren te beschrijven die van invloed kunnen zijn op sekseverschillen in schoolse keuzen.

Achtereenvolgens bespreken we theoretische referenties van de in het onderzoek opgenomen leerlingfactoren en gaan kort in op schoolfactoren.

Toekomstperspectief

De eerste invalshoek is die van de seksespecifieke toekomstperspectieven: een differentiële ontwikkeling van het toekomstperspectief van jongens en meisjes geeft richting aan differentiële vakkenkeuzen. Enerzijds gaat het om concrete plannen met betrekking tot vervolgopleiding en (de combinatie van) werk en gezin, anderzijds om attitudes in het onderwijs die samenhangen met die opties voor de toekomst. Bekend is dat de concrete toekomstplannen van jongens en meisjes nog steeds zeer verschillend zijn (o.a. Kuyper, 1990), aansluitend bij de huidige verschillen in arbeidsparticipatie (Niphuis-Nell, 1992; Dekkers, 1993a). De realistische of onrealistische toekomstverwachtingen van meisjes beïnvloeden hun belangstelling voor, prestaties in en keuze voor bepaalde opleidingen (Walberg, 1991). Deze verwachtingen kunnen de keuze beïnvloeden voor traditioneel vrouwelijke beroepen in bijvoorbeeld de zorgsector en het onderwijs, en daarmee samenhangend de keuze voor bepaalde vakken en opleidingen.

Attitudes

In samenhang met de toekomstplannen zijn aanzienlijke sekseverschillen aangetoond in met name het nut en de waarde die leerlingen toekennen aan vakken als wiskunde en natuur-

kunde (Eccles, Adler et al., 1985; Brush, 1985; Dekkers, 1990; Jörg & Man in 't Veld, 1987). Tegelijkertijd is aangetoond dat de perceptie door de leerling van het nut van wiskunde voor beide seksen de keuze voor en het succes in het vak wiskunde voorspelt (Armstrong, 1985; Ware & Dill, 1986).

Ook de interesse en het plezier in het vak hangen samen met keuze en succes (Murphy, 1978; Brush, 1985). Jongens vinden exacte vakken leuker dan meisjes (Armstrong, 1980), en meisjes die het vak leuker vinden brengen een cursus eerder tot een goed einde dan meisjes die het vak minder leuk vinden (Ware & Dill, 1986). In dit verband past echter enige terughoudendheid: een causale relatie met prestaties is nooit aangetoond. Het is evengoed denkbaar dat goede prestaties leiden tot positieve waardering van het vak.

Ten slotte kan het feit dat exacte vakken vaak nog als mannelijke vakken worden gestereotypeerd (Eccles et al., 1985; Zimmerer & Bennett, 1987) de keuze en prestaties van meisjes in die vakken beïnvloeden. Een aantal onderzoeksresultaten ontkracht echter deze veronderstelling: de stereotypering lijkt af te nemen (Zimmerer & Bennett, 1987), weinig samenhang met deelname te vertonen (Lantz & Smith, 1981) en sterker bij jongens dan bij meisjes te leven (Fennema & Sherman, 1978). Op de invloed van de mening van onder andere deze relevante anderen komen we in het volgende terug.

Attributie

Het zelfbeeld van leerlingen staat in relatie met schoolse prestaties en keuzen (Dekkers, 1990). Tegelijkertijd zijn er echter geen overtuigende onderzoeksresultaten die een verschil aangeven tussen jongens en meisjes in de mogelijkheid om een positief zelfconcept te ontwikkelen. Een uitzondering hierop vormt het zelfconcept met betrekking tot bekwaamheid in het vak wiskunde, waarop meisjes lager scoren. Het verschil tussen jongens en meisjes in zelfconcept is vooral in deelaspecten van het begrip onderzocht. De meeste onderzoeken passen binnen de attributie-theorie. Deze theorie richt de aandacht op de manier waarop personen een relatie leggen tussen oorzaken en gebeurtenissen en de daaruit voortkomende houdingen, verwachtingen en gedragingen bij volgende soortgelijke gebeurtenissen. Verondersteld wordt dat de invloed van de feitelijke prestaties op succesverwachtingen, resultaten en doelen in de toekomst geïntermedieerd wordt door causale attributies.

Weiner, Frieze et al. (1971) beschreven in hun (prestatie)motivatietheorie causale attributies van leerlingen met betrekking tot prestaties in het onderwijs. Succes of falen kan men toeschrijven aan diverse factoren zoals (gemis aan) bekwaamheid, (gebrek aan) inzet, (relatieve) moeilijkheidsgraad van de taak of aan (gebrek aan) geluk. Deze vier factoren kunnen vervolgens worden geclassificeerd langs twee dimensies: de 'Locus-of-Control'-dimensie, waarbij het van belang is of men prestaties toeschrijft aan factoren 'buiten zichzelf' (geluk, moeilijkheidsgraad) of factoren 'binnen zichzelf' (talent, inzet) en de stabiliteits-dimensie, waarbij het van belang is of men prestaties toeschrijft aan stabiele factoren (aanleg, taakmoeilijkheid) of aan onstabiele factoren (inzet, geluk). De resultaten van het onderzoek naar sekseverschillen in 'Locus of Control' (LOC) zijn niet eenduidig, maar in grote lijnen kan worden geconcludeerd dat sekseverschillen in attributie hierin bestaan dat meisjes hun falen in mathematische vakken wat meer toeschrijven aan gebrek aan bekwaamheid en jongens meer aan een gebrek aan inspanning (Eccles et al., 1985; Ware, Steckler & Leserman, 1985).

Gezinsachtergrond

De invloed van het opleidingsniveau en het beroep van beide ouders op de vakkenpakketkeuze van hun kinderen is vaak aangetoond. Geconcludeerd werd onder andere dat ouders van vrouwelijke 'science' studenten hoger opgeleid zijn dan die van mannelijke (Greenfield, Holloway & Remus, 1982; Jagacinski & Lebold, 1981), dat sociaal-economische achtergrond grote invloed heeft op niet traditionele keuzen (Bijsterveld, 1986) en dat leerlingen uit hogere milieus meer exacte vakken kiezen (Kodde & Theunissen, 1985; Thomas, 1984; Ware et al., 1985). Daarnaast werd aangetoond, dat niet zozeer het feit dat de moeder werkt van belang is voor technische keuzen van hun dochter (Van Vonderen & De Raaff, 1981), maar wel of dat een baan met aanzien is (Ware & Lee, 1985). Ten slotte lijkt een technisch beroep van de vader de dochter eerder op een

exact of technisch spoor te zetten (Van Vonderen & De Raaff, 1981; Granstam, 1985).

Relevante anderen

Behalve de indirecte invloed van het ouderlijk milieu kan ook (de inschatting van) het directe oordeel van de ouders over geschiktheid van bepaalde vakken een rol spelen bij de vakkenkeuze van de leerling. Ondersteuning van relevante anderen blijkt van invloed te zijn op de keuze van meisjes voor wiskunde (Lantz & Smith, 1981). Of de mening van de ouders terecht of onterecht is, deze is gerelateerd aan wat de leerling zelf denkt over zijn of haar bekwaamheden, aan zijn of haar verwachtingen over het vak en aan opleidingskeuzen (Yee, Jacobs & Goldsmith, 1986).

Ook de invloed van o.a leerkrachten is van belang. Deze invloed is in het voortgezet onderwijs aanzienlijk moeilijker aan te tonen dan in het basisonderwijs (Van der Hoeven & Jungbluth, 1988). Kuyper en Meulenbeld (1989) kunnen geen eenduidig effect van leerkrachten op de keuze voor wiskunde in het AVO aantonen. Over de invloed van leeftijdsgenoten en broers en zussen is weinig bekend. De mening van mannelijke leeftijdsgenoten lijkt belangrijk voor het gedrag (en dus ook de vakkenkeuzen) van meisjes in de tienerleeftijd (Lantz & Smith, 1981).

(Voorafgaande) leerprestaties

Over sekseverschillen in cognitieve bekwaamheden (ruimtelijk inzicht etc.), de oorzaken daarvan en de gevolgen voor (m.n. wiskunde) prestaties is veel gepubliceerd. We verwijzen hiervoor naar overzichtswerken zoals van Maccoby en Jacklin (1974) en Linn en Peterson (1985). De resultaten van de besproken onderzoeken spreken elkaar vaak tegen. Recente resultaten van meta-analyses en syntheses van studies van de afgelopen 20 jaar wijzen uit dat sekseverschillen in aanleg voor exacte en technische vakken zeer gering zijn (Linn & Hyde, 1989), en bovendien beïnvloedbaar: interventies zouden prestatieverschillen kunnen voorkomen.

Dat neemt niet weg dat er duidelijk sprake is van verschillen in prestaties in exacte vakken. Meisjes scoren systematisch lager dan jongens op wiskundetoetsen, bijvoorbeeld in de Scolastic Achievement Test (SAT) in de Verenigde Staten en op CITO-toetsen (o.a. Kremers, 1990) in Nederland. Hanna (in press) concludeert na heranalyse van de SIMS-data van de IEA-studies, dat in de meeste landen (12 van de 15) op 17-jarige leeftijd jongens beter zijn in wiskunde dan meisjes. Dat er een samenhang is tussen prestatiescores en keuzen is in Nederland onder andere aangetoond door Kuyper en Meulenbeld (1989) voor vakkenkeuze in het voortgezet onderwijs, en door Dekkers (1990) voor de overgang Voortgezet Onderwijs/ Technische Universiteit. In het laatste onderzoek moest de hypothese worden verworpen dat er op het eind van het Voortgezet Onderwijs sprake is van een reservegroep meisjes die op basis van hun prestaties evengoed als jongens voor een technische universiteit hadden kunnen kiezen. De meisjes bleken minder geschikte exacte vakken te hebben gekozen en bovendien minder goed gepresteerd te hebben in deze vakken.

Hoewel er op de diverse keuzemomenten sprake is van een samenhang tussen prestaties en keuzen, moet tegelijkertijd worden vastgesteld dat de zeer grote sekseverschillen op het moment van de vakkenpakketkeuze geenszins de tussen de seksen weinig uiteenlopende leerprestaties aan het eind van het basisonderwijs reflecteren.

Schoolfactoren

Het feit of scholen verschillen in seksespecifieke keuzen en prestaties van hun leerlingen is nauwelijks gedocumenteerd. Wel is veelvuldig gezocht naar schoolse factoren die van invloed kunnen zijn op exacte en technische keuzen van meisjes. Oakes (1990) beschrijft de belangrijkste schoolse factoren waarover in de Angelsaksische literatuur verslag is gedaan. Ze concludeert dat factoren als 'curriculum tracking', keuzebegeleiding en advies en ervaringen in de klas soms seksespecifiek zijn, maar wijst er tegelijkertijd op dat deze nog niet de oorzaak behoeven te zijn van verschillen in deelname aan wiskunde (Oakes, 1990). Schoolfactoren zijn bovendien schoolsysteemgebonden en moeten deels in nationaal onderzoek worden bestudeerd. In Nederland heeft het onderzoek zich tot nu toe toegespitst op de invloed van decanen en keuzebegeleiding (Dekkers, 1985), op de invloed van leerkrachten (Kuyper & Van der Werf, 1987), op

kenmerken van het aangeboden curriculum (Jörg & Man in 't Veld, 1987) en op formele schoolkenmerken en schoolbeleid (Dekkers, 1985; Bosker, 1990; Kristensen & Jenneskens, 1991). Uit (voornamelijk outlier)studies komt naar voren dat op sommige scholen een meisje op de drie jongens wiskunde kiest, terwijl op andere scholen de ratio bijna een is. Een belangrijke conclusie van dit soort onderzoek is dat specifieke op de emancipatie van meisjes gerichte maatregelen minder samenhangen met een hoge wiskunde populariteit dan met een goed georganiseerde niet-specifieke keuzebegeleiding; adviezen moeten meer gebaseerd zijn op de prestaties van individuele meisjes dan op hun attitudes (Kristensen & Jenneskens, 1991). Een andere bevinding is dat schoolfactoren die van invloed lijken te zijn, zoals een doelgericht schoolbeleid en een coherent leerkrachtteam, alleen keuzeverschillen tussen jongens verklaren (Bosker, 1990).

Op schoolverschillen en schoolfactoren gaan we in dit artikel niet nader in; in andere publikaties wordt hierover verslag gedaan (o.a. Bosker & Dekkers, 1993a en b).

2 Vraagstelling

De vraagstelling voor dit artikel heeft betrekking op leerlingfactoren die van invloed zijn op sekseverschillen in keuzen voor (exacte) vakken. Het gaat om de relatieve invloed van leerlingvariabelen als gezinsachtergrond, toekomstperspectief, leerprestaties, attitudes, attributies en de invloed van relevante anderen op vakkenkeuze; daarnaast gaat het om ontwikkelingen in de onafhankelijke en afhankelijke variabelen, met name in de tijd tussen eind basisschool en het moment van vakkenkeuze in het voortgezet onderwijs. Meer specifiek is de vraag in welke mate wijzigingen in de onafhankelijke variabelen samenhangen met wijzigingen in (voorgenomen) vakkenkeuzen, en of deze samenhangen en ontwikkelingen daarin seksespecifiek zijn. Het gaat om de exacte vakken wiskunde en natuurkunde, en om het vergelijkingsvak Frans. We bespreken resultaten voor HAVO- en VWO-leerlingen.

3 Onderzoeksdesign

Niet alleen de aard van de variabelen, maar ook de optiek die ten grondslag ligt aan het hele onderzoek, namelijk dat er sprake is van een ontwikkelingsproces dat leidt tot seksespecifiek keuzegedrag, vraagt om een longitudinale onderzoeksopzet.

In de longitudinale studie wordt gedurende een aantal jaren een groep leerlingen gevolgd, vanaf de intrede in het voortgezet onderwijs tot aan de definitieve vakkenpakketkeuze (voor MAVO eind leerjaar 2, voor HAVO eind leerjaar 3, voor VWO eind leerjaar 4). Het betreft 2500 leerlingen van 25 aselect gekozen scholengemeenschappen voor AVO/VWO die in het schooljaar 1989/1990 in het eerste brugjaar zaten. Het eerste jaar ging het om vier brugklassen per school, elk met ongeveer 25 leerlingen. Alleen de leerlingen die overgingen zijn de daaropvolgende jaren gevolgd. Uitval in de vorm van zittenblijven, maar ook door ziekte of verzuim van leerlingen en door uitval van een hele school resulteerde het tweede jaar in een onderzoeksgroep van ongeveer 1850 leerlingen. Na het tweede jaar vielen ook de leerlingen die ondertussen MAVO hadden gekozen af, en resteerde er uiteindelijk een groep van 900 leerlingen (helft HAVO, helft VWO) die alle drie de jaren aan alle metingen hadden deelgenomen. In het VWO moet in het vierde jaar nog een nameting plaatsvinden om de definitieve vakkenkeuze vast te stellen.

4 Instrumenten

Elk jaar is aan het eind van het schooljaar bij alle leerlingen een wiskundetoets en een vragenlijst afgenomen.

De *wiskundetoets* bestond de eerste twee jaar uit een gedeelte van de wiskundetoets voor de brugklas van het CITO (twee jaar identiek), het derde jaar uit de meest valide helft van een wiskundetoets die door het CITO is ontwikkeld in het kader van de evaluatie van de basisvorming. De toetsen namen steeds een lesuur in beslag. De beide toetsen zijn zo samengesteld dat vergelijkingen mogelijk zijn van de prestaties in de opeenvolgende jaren.

Naast de wiskundetoets hebben de leerlingen elk jaar een *vragenlijst* ingevuld die de instrumentatie bevatte van de hierboven besproken leerlingvariabelen, nl.:

Vakkenkeuze (afhankelijke variabele)

De (voorgenomen) keuze van wiskunde, natuurkunde en Frans is opgenomen in een vraag waarin de leerling van alle mogelijke keuzevakken moest aangeven of hij/zij deze zeker, waarschijnlijk wel, waarschijnlijk niet, zeker niet zal kiezen, of het nog niet weet. Voor HAVO ging het om wiskunde algemeen (het onderzoek vond grotendeels plaats vóór de invoering van wiskunde A en B op alle HAVO's), voor VWO om wiskunde A en B.

Toekomstperspectief

De concrete toekomstplannen van de leerlingen zijn enerzijds gemeten aan de hand van vragen over het niveau en de sector van de geplande vervolgopleiding, anderzijds aan de hand van een aantal vragen over het aantal dagen dat men van plan is te gaan werken, in combinatie met vragen over toekomstig samenwonen/trouwen en kinderen. De vragen zijn grotendeels samengesteld op basis van eerder onderzoek op dit terrein (Kuyper & Meulenbeld, 1989; Jörg & Man in 't Veld, 1987; Dekkers, 1990).

Attitudes

Het oordeel over het vak wiskunde (en Frans) is gemeten aan de hand van een aantal items die relevante schalen bleken te vertegenwoordigen in eerder onderzoek naar de keuze van wiskunde bij MAVO, HAVO en VWO leerlingen (Kuyper & Meulenbeld, 1989). In negen sterk gebleken bipolaire negen-puntsitems en twee vijfpuntsitems per vak werd een oordeel gevraagd over het plezier in het vak, de moeilijkheid, het nut, seksegeschiktheid, en assertiviteit met betrekking tot het vak.

Attributie

De leerlingen is gevraagd voor het vak wiskunde (en Frans) aan te geven waaraan zij een lager cijfer dan gebruikelijk zouden toeschrijven, en aan welke reden een hoger cijfer. Steeds moesten 100 punten over vier redenen worden verdeeld (ik ben eigenlijk minder goed, ik heb niet hard genoeg gewerkt, het vak is toch wel moeilijk, toevallig pech; of andersom). Deze instrumentatie bleek in eerdere onderzoeken (Kuyper & Meulenbeld, 1989) de beste manier om attributies in dit kader te operationaliseren. Daarnaast is ter controle voor het vak wiskunde nog een vraag met anders geformuleerde redenen voor succes en falen voorgelegd (2 x 8 items).

Gezinsachtergrond

Met betrekking tot de gezinsachtergrond van de leerlingen is gevraagd naar de gezinssamenstelling, woonsituatie, opleiding en beroep van beide ouders, regio en herkomstland. De instrumentatie is grotendeels ontleend aan onderzoek op het terrein van schoolloopbaan en milieu.

Relevante anderen

Deze variabele betreft het door de leerling gepercipieerde advies van relevante personen uit zijn omgeving over de keuze van wiskunde (en Frans). Aan de leerlingen is gevraagd op een 7-puntsschaal aan te geven hoe waarschijnlijk het is dat de betreffende persoon wiskunde (of Frans) een geschikt eindexamenvak voor hen vindt.

Voorafgaande leerprestaties

Behalve de score op de CITO-wiskundetoetsen zijn van de leerlingen ook rapportcijfers bekend. Het eerste jaar zijn deze bij de scholen opgevraagd, het tweede en derde jaar aan de leerlingen zelf. Het gaat om cijfers voor wiskunde, natuurkunde, Engels en Frans.

5 Analyse

De hoofdvraag in de analyses is of verandering in motivatie etc. samenhangt met verandering in keuze, ofwel: als een leerling meer in motivatie verandert dan een andere leerling, verandert deze dan ook meer in keuze dan de andere leerling. Het betreft dus een zgn. 'parameter of change' benadering. Deze vraag is beantwoord door correlatie-/multiple regressie analyses met de verandering in keuze als afhankelijke variabele en de verandering in motivatie etc. als onafhankelijke variabele. Als veranderingsmaat is de verschilscore tussen meting 3 en meting 1 genomen.

De gevolgde analysemethode lijkt misschien op het eerste gezicht niet de meest adequate voor analyse van longitudinale data. Daarom noemen we kort een aantal voor de hand liggende alternatieven en geven we aan waarom die niet geëigend zijn om de onderzoeksvraag te beantwoorden. In de eindrapportage van het onderhavige onderzoek worden de alternatieven uitvoerig besproken (Dekkers, 1993b).

Een alternatief lijkt op het eerste gezicht de 'change in parameter' benadering. Hierin wordt, bijvoorbeeld met een LISREL procedure, bekeken/getoetst of de samenhang tussen bijvoorbeeld motivatie en keuze over de metingen heen verandert. Motivatie en keuze kunnen daarbij latente variabelen zijn die ieder door een aantal gemeten variabelen worden geïndiceerd. Een verandering (of ontbreken van een verandering) in de samenhang tussen motivatie en keuze over de tijd, zegt echter niets over de samenhang van de verandering in motivatie op de verandering in keuze, omdat de structuur binnen metingen heel anders kan zijn dan de structuur van de veranderingen.

Als volgend alternatief zou men zich een multilevel analyse kunnen voorstellen met de drie metingen als lager niveau en leerlingen als hoger niveau. Deze analyse geeft o.a. de gemiddelde binnen-persoons samenhang tussen motivatie en keuze. Deze samenhang geeft ook geen inzicht in de vraag of leerlingen die meer in motivatie veranderen ook meer in keuze veranderen.

Ook een multilevel analyse met op het lager niveau als predictoren: motivatie, meettijdstippen en het produkt van deze beide, geeft niet wat we zoeken; evenmin een multilevel analyse met alleen meettijdstip als predictor op het lager niveau. Hoewel op zich interessante, en veel gebruikte, analysemethoden voor longitudinale data, beantwoorden deze ook niet onze vraag of een verandering in keuze samengaat met een verandering in motivatie.

Ook een 'klassieke' analyse als ANOVA voor herhaalde metingen met geslacht resp. schooltype en meting als factoren geeft alleen het verloop van de gemiddelde keuze over de tijd, uitgesplitst naar geslacht resp. schooltype, inclusief interactie toetsing. In een MANOVA voor herhaalde metingen wordt het gemiddelde van contrasten in de keuze geanalyseerd, bijvoorbeeld meting 1 ten opzichte van het gemiddelde van meting 2 en 3. Ook deze analyses beantwoorden onze onderzoeksvraag dus niet.

In de door ons toegepaste analyse van de samenhang van veranderingen is gebruik gemaakt van de verschilscore tussen meting 3 en meting 1 als veranderingsmaat. De redenering hierachter is als volgt: voor drie metingen met een gelijk tijdsinterval tussen de metingen is de regressiecoëfficiënt door de drie metingen proportioneel aan de verschilscore tussen meting 3 en meting 1, nl. de regressiecoëfficiënt = (verschil 3-1)/2. Als maat voor de verandering over de drie metingen speelt de waarde op de tweede meting dus geen rol.

Alhoewel soms getwijfeld wordt aan het gebruik van verschilscores, o.a. op grond van vermeende onbetrouwbaarheid, toont Willett (1988/89) duidelijk aan dat ook een verschilscore tussen twee metingen een niet-vertekende schatting is van de ware verandering over de tijd, en daarom als analysemateriaal niet onderdoet voor een enkele meting van een kenmerk.

Voor onze data zou men verder kunnen tegenwerpen dat van de tweede meting geen gebruik wordt gemaakt en de analyse dus 'inadequaat' is. Dit argument is onterecht als het gaat om de bepaling van de regressiecoëfficiënt door de drie metingen nl. bij *3* metingen met gelijke tijdsintervallen speelt de tweede meting geen rol bij de berekening van de regressiecoëfficiënt. De tegenwerping is terecht als wordt beweerd dat de tweede meting gebruikt zou kunnen worden voor de bepaling van de betrouwbaarheid van de regressiecoëfficiënt door de drie metingen. Deze betrouwbaarheid zou dan gebruikt kunnen worden voor de bepaling van de 'ware waarde' van de regressiecoëfficiënt, die vervolgens in de analyse gebruikt zou kunnen worden. Deze ware waarde kan bepaald worden met multilevel analyse, echter alleen per analyse variabele apart. De betrouwbaarheidscoëfficient, die gebruikt wordt voor de zgn. Bayesiaanse schatting van de regressiecoëfficiënt per leerling blijkt echter voor alle leerlingen hetzelfde te zijn (zie ook Dekkers, 1993b). Voor verschillende variabelen zijn deze betrouwbaarheids-coëfficiënten wel weer anders, maar tussen leerlingen gelijk. Omdat we omwille van de vergelijkbaarheid tussen de parameterschattingen uit zijn gegaan van gestandaardiseerde veranderingscores,

kunnen deze verschillende betrouwbaarheidscoëfficiënten tussen variabelen geen invloed hebben op de parameter schattingen tussen veranderingsmaten. Om deze reden hebben we dan ook gebruik gemaakt van de ongecorrigeerde veranderingsmaten.

Algemeen kunnen we zeggen dat het probleem bij een 'parameter of change' analyse via multilevel analyse is dat de veranderingen, inclusief betrouwbaarheidscorrectie, voor zowel afhankelijke als onafhankelijke variabelen *gelijktijdig* moeten worden geanalyseerd, een procedure die we nog niet zijn tegengekomen.

In de analyses is verder gezocht naar sekse- respectievelijk schooltype-specifieke verschillen; omdat de mogelijkheid bestaat dat de samenhang tussen de veranderingen sekse- of schooltype specifiek zijn, zijn de analyses apart voor jongens en meisjes en voor HAVO en VWO uitgevoerd.

Voor afzonderlijke analyses binnen de meisjes-, jongens-, HAVO- en VWO-groep is het totale aantal indicatoren in het model (ongeveer 60, zie instrumentatie) wat te groot in relatie tot het aantal respondenten binnen elk van deze groepen (ruim 400). Dit zou tot instabiele schattingen van de directe werkingen van de indicatoren kunnen leiden. Ook zijn de correlaties tussen verschillende predictoren erg hoog (m.n. tussen de items in het blok 'geschiktheid volgens relevante personen'). Deze multicollineariteit leidt ook tot onbetrouwbare en instabiele schattingen. Daarom is eerst via factoranalyse gekeken of de samenhangen tussen de verschilscores zodanig gestructureerd zijn dat tot een zinvolle indikking van indicatoren gekomen kan worden.

Het volgende overzicht geeft de variabelen, oorspronkelijke indicatoren en factoren/(rest) items weer.

Variabelen	*Indicatoren*	*Factoren/items*
Toekomstperspectief	– Interesse *vervolgopleiding* (15 opl.)	– 3 factoren en 2 aparte items
	– Geplande aantal dagen werken (met kinderen)	– 2 items
Attitudes	– Oordeel over wiskunde en Frans (9 + 2 items per vak)	– 3 factoren en 3 aparte items (wiskunde) – 2 factoren en 2 items (Frans)
Attributies	– Reden hoger cijfer 4 items per vak Reden lager cijfer 4 items per vak Voor wiskunde extra 2 × 8 items	– 6 factoren + 2 items (wiskunde) – voor Frans niet opgenomen
Gezinsachtergrond	– Opleiding vader	– 1 score
Relevante anderen	– Oordeel 8 relevante personen over wiskunde en Frans	– 1 factor (voor wiskunde en Frans apart)
Leerprestaties	– CITO-toets	– 1 score
	– Rapportcijfers (wiskunde, natuurkunde en Frans)	– 1 score per vak
Vakkenkeuze	– Voor HAVO → wiskunde Voor VWO → wiskunde B, verder Frans en natuurkunde	– 1 score per vak

Inhoudelijk betreft het de volgende factoren:

- toekomstperspectief (vervolgopleiding):
  - interesse talen/toerisme/sociaal
  - interesse verzorging
  - interesse techniek/exact/informatica

  en economie en onderwijs apart.
- attitude t.o.v. wiskunde:
  - wiskunde is leuk (makkelijk, interessant)
  - wiskunde is nodig
  - wiskunde is geschikt

  en drie aparte items.
- attitude t.o.v. Frans
  - Frans is makkelijk, leuk, nodig, interessant
  - Frans is geschikt

- attributie wiskunde: ander cijfer dan normaal ...
  - ligt aan mezelf
  - ligt aan omstandigheden
  - zegt niets over een volgende keer
  - heeft altijd dezelfde oorzaak
  - ligt aan oorzaken buiten mij
  - ligt steeds aan andere dingen

  en twee aparte items.

- relevante personen:
  de items in dit blok correleren hoog genoeg om in feite één factor te vormen. Een 3-factor oplossing geeft wel enige differentiatie, nl. in ouders/broer, vriend/school en zus, maar deze factoren correleren onderling .50-.60. Dit is te hoog voor predictoren in een multiple regressie; effecten kunnen ontstaan (en blijken ook te ontstaan). Daarom is in de analyse gebruik gemaakt van de 1-factor oplossing:
  - wiskunde/Frans geschikt volgens relevante personen.

De in het model verder opgenomen indicatoren opleiding vader, aantal dagen werken in combinatie met gezin, scores wiskundetoets en rapportcijfers zijn niet ingedikt. Voor de te verklaren variabele 'keuze wiskunde' ten slotte is voor HAVO wiskunde en voor VWO wiskunde B opgenomen.

6 Resultaten

Om vast te stellen of verandering in een predictor samengaat met verandering in de (voorgenomen) keuze voor wiskunde, natuurkunde en Frans, is het zinvol eerst naar de totale samenhang van de (verandering in) predictor met de (verandering in) keuze te kijken. Omdat sekserespectievelijk schooltype-specifieke verschillen hier juist interessant zijn, kunnen deze samenhangen het beste apart voor jongens/meisjes respectievelijk HAVO/VWO bekeken worden. Voor de overzichtelijkheid zijn alleen die samenhangen in een tabel gezet waarbij binnen sekse en binnen schooltype minstens één van de twee samenhangen groter of gelijk aan .20 is.

De gemiddelde verandering in keuze inclusief de standaarddeviatie is steeds eerst vermeld (de schaal voor verandering in keuze loopt steeds van -4 tot +4; een negatieve verschuiving betekent een sterkere (voorgenomen) keuze).

Bevindingen voor wiskunde (Bevinding 1):

	M	J	HAVO	VWO
Wiskunde keuze:				
Gemiddelde verandering	.0	-.3	-.3	.0
(standaarddeviatie	1.7	1.5	1.6	1.6)

Gemiddeld wordt er dus het derde jaar iets zekerder voor wiskunde gekozen bij jongens en HAVO-leerlingen (Bevinding 1). Gezien de breedte van de veranderingschaal (-4 tot +4) zijn de gemiddelde verschuivingen gering. De variatie in de veranderingen is echter voldoende groot om een analyse van samenhangen tussen veranderingen zinvol te maken.

Tabel 2
Samenhang verandering in predictor met verandering in (voorgenomen) keuze wiskunde (Pearson r)

	M	J	HAVO	VWO
Hoog cijfer wiskunde	-.21	-.22	-.21	-.20
Wiskunde leuk	.31	.31	.35	.28
Wiskunde nodig			.22	.11
Geschikt vlg. rel. pers.	-.37	-.28	-.47	-.20
Interesse technische/ exacte opl.	.20	.14	.15	.20

De verandering in keuze wiskunde blijkt voornamelijk samen te gaan met een verandering in
- de (gepercipieerde) geschiktheid volgens relevante anderen
- het wiskunde leuk, interessant en makkelijk vinden

en wel zo dat een grotere geschiktheid volgens relevante personen en meer interesse/gemak in wiskunde samengaan met een zekerder keuze voor wiskunde (Tabel 2).

In mindere mate is er samenhang met de verandering in wiskunde cijfer, het wiskunde nodig vinden voor de toekomst (alleen bij HAVO-leerlingen), en een technische/exacte vervolgopleiding interessant vinden. Veranderingen in overige interesse-, attitude- en attributie-factoren, alsook in scores op de wiskundetoets en op arbeidsdeelnameplannen hangen voor geen enkele groep relevant samen

met veranderingen in wiskunde keuze.

Er is weinig verschil tussen jongens en meisjes. Ook is er meestal alleen een gradueel verschil tussen HAVO en VWO, en wel zo dat, (afgezien van een verandering in cijfer en een technische/wiskunde opleiding interessant vinden), verandering in keuze voor wiskunde bij VWO-leerlingen minder samengaat met verandering in de overige aspecten dan bij HAVO-leerlingen.

Bevindingen voor natuurkunde (Bevinding 2) (de predictoren zijn dezelfde als bij wiskunde):

	M	J	HAVO	VWO
Natuurkunde keuze:				
Gemiddelde verandering	.8	-.3	.8	-.2
(standaarddeviatie	1.9	1.9	2.0	1.9)

Gemiddeld heeft er bij meisjes en bij HAVO leerlingen een vermindering van keuze voor natuurkunde plaatsgevonden.

Tabel 3
Samenhang verandering in predictor met verandering in (voorgenomen) keuze natuurkunde

	M	J	HAVO	VWO
Hoog cijfer wiskunde			-.21	-.12
Wiskunde interessant	.21	.21	.29	15
Geschikt vlg. rel. pers.	-.20	-.16	-.21	-.15
Interesse technische/ exacte opl.	.30	.30	.32	.31

In tegenstelling tot de keuze voor wiskunde gaat de verandering in keuze voor natuurkunde het sterkst samen met het interessanter zijn gaan vinden van een vervolgopleiding techniek/informatica/wiskunde. Dit geldt voor jongens, meisjes, HAVO en VWO in gelijke mate (Tabel 3).

De overige samenhangen zijn overeenkomstig wat bij wiskunde keuze werd gevonden: bij VWO-leerlingen hangt de verandering in natuurkunde keuze minder sterk samen met de andere (op wiskunde betrekking hebbende) aspecten dan bij HAVO-leerlingen.

Bevindingen voor Frans (Bevinding 3):

	M	J	HAVO	VWO
Frans keuze:				
Gemiddelde verandering	.7	1.1	1.2	.5
(standaarddeviatie	1.7	1.9	1.9	1.7)

De keuze voor Frans is over het algemeen redelijk afgenomen respectievelijk minder zeker geworden, met name bij jongens en HAVO-leerlingen (Bevinding 3).

Tabel 4
Samenhang verandering in predictor met verandering in (voorgenomen) keuze Frans

	M	J	HAVO	VWO
Hoog cijfer Frans			-.05	-.21
Frans leuk, interessant	.51	.37	.52	.41
Geschikt vlg. rel. pers.	-.45	-.40	-.50	-.35
Interesse technische/ exacte opl.	.22	.09	.20	.12

Een verandering in de keuze voor Frans hangt het sterkst samen met verandering in het Frans leuk en interessant vinden en met verandering in de ingeschatte geschiktheid volgens relevante personen (Tabel 4). Verder gaat bij meisjes en HAVO-leerlingen de verandering in keuze Frans samen met 'interesse voor vervolgopleiding techniek/wiskunde', en wel zo dat een relatief zekerder keuze voor Frans samengaat met een relatief *on*interessanter zijn gaan vinden van een technische/exacte vervolgopleiding.

De hierboven gepresenteerde correlaties zijn de som van een directe werking en een indirecte werking via samenhang met de andere variabelen. Om de directe werking van de variabelen af te splitsen zijn multiple regressies uitgevoerd. Tevens is een eenvoudig verklaringsmodel opgelegd dat bestaat uit 4 blokken variabelen (Model I).

De blokken worden verondersteld in deze volgorde een causale werking te hebben. Door een opeenvolging van multiple regressies (nl. blok 1+2+3 op blok 4, blok 1+2 op blok 3, blok 1 op blok 2) zijn de verschillende padcoëfficiënten (eigenstandige werkingen) berekend.

Bij de analyse van meisjes resp. jongens is schooltype vooraf uitgepartialiseerd om te voorkomen dat tussen-schooltype verschillen in gemiddelden de correlaties beïnvloeden. Bij de analyse van HAVO resp. VWO is om dezelfde reden sekse vooraf uitgepartialiseerd.

Als we ook hier weer alleen die paden opnemen die een lading hebben van minstens .20 in een van de groepen meisjes/jongens resp. HAVO/VWO, dan blijkt dat het uiteindelijke model voor de 4 groepen hetzelfde is.

Model I

Verklaringsmodel

voor wiskunde en natuurkunde:

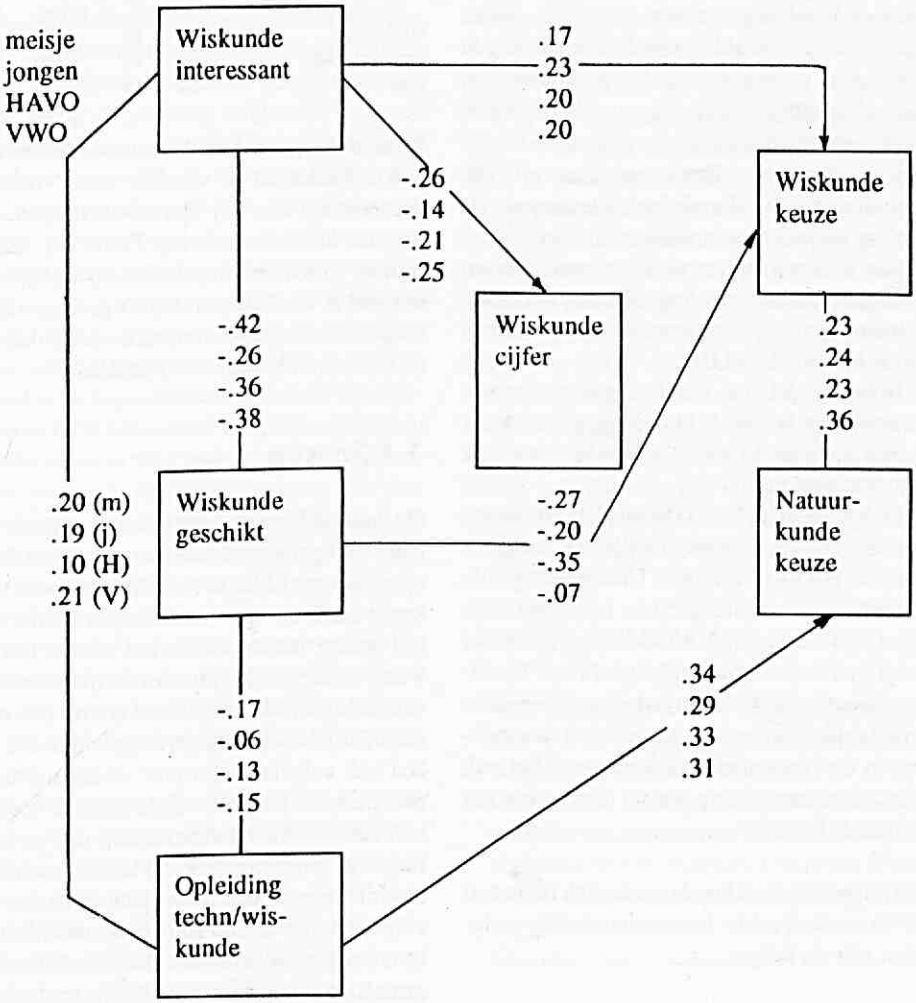
1	2	3	4
	Oordeel wiskunde		
	Attributie wiskunde	CITO-score	Wiskunde keuze
Opleiding vader	Relevante personen	Cijfer wiskunde	Natuurkunde keuze
	Interesse vervolgopl.		
	Later werken		

voor Frans:

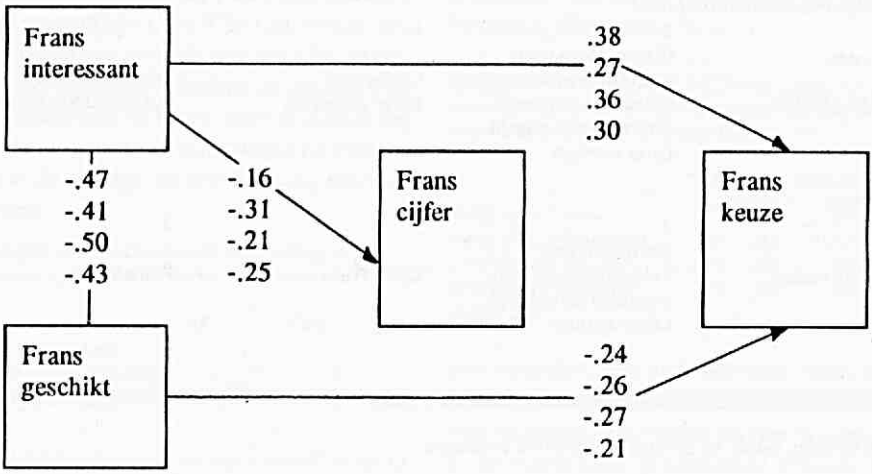
1	2	3	4
	Oordeel Frans		
Opleiding vader	Relevante personen	Cijfer Frans	Frans keuze
	Interesse vervolgopl.		
	Later werken		

Model II

Exacte keuzen, sekse- en schooltypespecifiek verklaard



Model III
Keuze Frans, sekse- en schooltypespecifiek verklaard



Voor wiskunde/natuurkunde geldt: er is sprake van een directe werking van de verandering in 'wiskunde interessant/makkelijk vinden' en van 'wiskunde geschikt volgens relevante personen' op de verandering in wiskunde keuze. Ook gaat er een directe werking uit van verandering in 'wiskunde leuk/interessant vinden' op verandering in wiskunde cijfer.
Verder is er een directe werking van de verandering in 'vervolgopleiding techniek/wiskunde is interessant' op de verandering in natuurkunde keuze (Model II).

In het model zijn ook de correlaties tussen variabelen in hetzelfde blok aangegeven zodat te zien is via welke paden er sprake is van een indirecte werking.

Ook bij het vak Frans (Model III) zijn de samenhangen voor VWO-leerlingen zwakker dan voor HAVO-leerlingen. Uitzondering is de verandering in de hoogte van het cijfer voor Frans, dat alleen bij VWO-leerlingen samenhangt met verandering in keuze Frans. Overigens hangt verandering in interesse in bepaalde vervolgopleidingen niet samen met verandering in de (voorgenomen) keuze van het vak Frans; deze samenhang bestaat alleen voor natuurkunde keuze.

De (proporties verklaarde varianties (adjusted R**2) voor de drie keuzeverandering-variabelen zijn als volgt:

	M	J	HAVO	VWO
Wiskunde	.17	.13	.24	.10
Natuurkunde	.11	.14	.14	.10
Frans	.32	.19	.32	.19

Te zien is dat ook hier de keuzeverandering bij VWO-leerlingen in mindere mate voorspeld kan worden dan bij HAVO-leerlingen. Ook kan de keuzeverandering Frans bij jongens minder worden voorspeld dan bij meisjes. Opvallend is verder de (relatief) geringe verklaringskracht van de betreffende variabelen voor de natuurkunde keuze van meisjes.

7 Discussie

De vraagstelling waarop de bovenstaande analyses zijn gebaseerd had met name betrekking op sekseverschillen in (verklaringen voor) vakkenkeuzen, en op veranderingen hierin in de tijd tussen het eerste en het vierde leerjaar. Voor jongens en meisjes blijkt eenzelfde verklaringsmodel te gelden: vooral (veranderingen in) de gepercipieerde geschiktheid voor een vak volgens 'relevante' anderen (ouders, docenten etc.) en een vak leuk, interessant en makkelijk vinden hangen samen met veranderingen in de (voorgenomen) keuze van een vak (en interesse in een technische of exacte vervolgopleiding met de keuze van natuurkunde). Er is dus geen sprake van sekseverschillen in de samenhang tussen predictoren en vakkenkeuze. Dat wil niet zeggen dat er geen sekse-

verschillen in de afzonderlijke variabelen bestaan. Eerder is gerapporteerd dat er sekseverschillen in vakkenkeuze bestaan (met name voor wiskunde B, natuurkunde en Frans); ook zijn er sekseverschillen in de relevante predictoren (Dekkers, 1992a). De samenhangen zien er dus wel voor jongens en meisjes hetzelfde uit, ze liggen echter op een ander scoreniveau; jongens scoren hoger op (de predictoren van) wiskunde en natuurkunde, meisjes op (die van) Frans. Dit geldt voor alle variabelen die een rol spelen in het model.

Er is daarnaast ook sprake van sekseverschillen op predictoren die geen samenhang vertonen met vakkenkeuze. Met name voor toekomstplannen (werk, gezin) geldt dat er grote (traditionele) sekseverschillen bestaan (Dekkers, 1992b). Deze plannen blijken echter noch voor jongens, noch voor meisjes repercussies te hebben voor het al dan niet kiezen van (exacte) vakken; ze passen niet in het verklaringsmodel. Hetzelfde geldt voor wiskunde nodig of geschikt vinden: er zijn wel sekseverschillen, echter de invloed van deze factoren in het model is te klein om de door ons gestelde relevantiegrens te passeren. Ook de attributievariabelen spelen geen rol in het model; wel verklaren meisjes hun falen in het vak wiskunde vaker door de moeilijkheid ervan, en jongens door hun luiheid; ze kiezen het vak er echter niet meer of minder om. Het meest opvallend is de lage samenhang tussen veranderingen in CITO-scores en wiskundecijfers met veranderingen in wiskunde keuze. De rapportcijfers wiskunde vertonen overigens een grotere samenhang met wiskunde keuze dan de (meer leerkracht-onafhankelijke) toetsscores; in het model valt deze echter weg door de hoge samenhang van cijfers met het vak leuk en interessant vinden. De opleiding van de ouders ten slotte verschilt niet voor de seksen, maar speelt tegen de verwachting in noch voor jongens noch voor meisjes, een relevante rol bij hun vakkenkeuze.

Dat het model/de variabelen meer verklaringskracht heeft/hebben voor HAVO dan voor VWO kan niet al te definitief worden geïnterpreteerd. In de analyses is de definitieve keuze van VWO-ers nog niet opgenomen (eind 4e leerjaar), terwijl het aannemelijk is dat daarin het laatste jaar nog aanzienlijke veranderingen optreden.

Een opmerking dient ten slotte gemaakt te worden over de recursiviteit van het model; de verbanden kunnen ook andersom worden geïnterpreteerd. Een goed wiskunde cijfer halen kan tot gevolg hebben dat het vak leuk en interessant wordt gevonden, een vak per se willen volgen kan tot betere cijfers leiden en een geplande keuze kan anderen tot een positief oordeel over de geschiktheid van het vak leiden. Vandaar een kanttekening bij de volgorde van de variabelen in het model; deze lijkt voorlopig in tijd de meest logische, en wordt bovendien in de literatuur op dezelfde manier gehanteerd, maar is niet noodzakelijk.

Terugkerend naar het maatschappelijk probleem van sekse-ongelijke keuzen in het onderwijs, waarmee vooral de negatieve keuze van meisjes voor exacte vakken en de daarmee samenhangende ongelijke arbeidsmarktkansen worden bedoeld, kan worden geconcludeerd dat een gelijktijdige aanpak van relevante variabelen tot de beste resultaten zal leiden. Zowel bepaalde attitudes van de leerling (plezier, interesse) en van relevante anderen (geschiktheid) ten opzichte van de betreffende vakken, als de prestaties daarin zijn aspecten waarop naast elkaar de aandacht moet worden gericht; minder zinvol lijkt directe aandacht voor toekomstige arbeids- en gezinsplanning of voor het nut voor een vak voor de toekomst. Omdat de richting van de samenhangen niet eenduidig is lijkt het daarnaast zinvol niet-traditionele vakken- en opleidingskeuzen op zich onder de aandacht van de leerlingen te brengen. Het resultaat kan dan zijn dat het scoreniveau van meisjes op de predictoren én op de afhankelijke variabele dat van de jongens benadert.

Literatuur

Armstrong, J. M. (1980). *Achievement and participation in mathematics: an overview.* Washington DC: National Institute of Education.

Armstrong, J. M. (1985). A national assessment of participation and achievement of women in mathematics. In S. F. Chipman, L. R. Brush & D. M. Wilson (Eds.), *Women and mathematics: Balancing the equation* (pp. 59-94). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bijsterveld, A. (1986). *Tussen traditie en emancipatie. Onderzoek naar de verschillen tussen traditionele en niet-traditionele meisjes op het VWO.* Groningen: RUG.

Bosker, R. J. (1990). *Extra kansen dankzij de school?* Nijmegen: ITS/OOMO-reeks.

Bosker, R. J., & Dekkers, H. (1993a). *School differences in producing gender related subject choices.* Paper ICSEI 1993, Norköping.

Bosker, R. J., & Dekkers, H. (1993b). *Schoolverschillen in seksespecifieke vakkenkeuzen.* Paper Onderwijsresearchdagen 1993, Maastricht.

Brush, L. R. (1985). Cognitive and affective determinants of course preference and plans. In S. F. Chipman, L. R. Brush & D. M. Wilson (Eds.), *Women and mathematics: Balancing the equation* (pp. 123-150). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dekkers, H. (1985). *Soms kiezen meisjes anders.* Nijmegen: ITS.

Dekkers, H. (1990). *Seksespecifieke studiekeuzen in het wetenschappelijk onderwijs.* Nijmegen: ITS.

Dekkers, H. (1992a). *Determinanten van seksespecifieke vakkenkeuzen in het Voortgezet Onderwijs.* Reader met tussentijdse rapportage. Nijmegen: ITS.

Dekkers, H. (1992b). *Determinants of sex-specific subject choices in secondary education.* Paper Onderwijsresearchdagen 1992, Enschede.

Dekkers, H. (1993a). *Onderwijsemancipatiebeleid en economische zelfstandigheid van vrouwen in Noordwest Europa.* Nijmegen: ITS.

Dekkers, H. (1993b). *Determinanten van seksespecifieke vakkenkeuzen in het Voortgezet Onderwijs.* Reader, deel 2. Nijmegen: ITS.

Dossey, J. A., Mullis, I. V. S., Lindquist, M. M., & Chambers, D. L. (1988). *The mathematics report card. Are we measuring up?* Princeton, NJ: Educational Testing Service.

Eccles, J. Adler, T. F., Futterman, R., Goff, S. B., Kaxzala, C. M., Meece, J. L., & Midgley, C. (1985). Self-perceptions, task perceptions, socializing influences, and the decision to enroll in mathematics. In S. F. Chipman, L. R. Brush & D. M. Wilson (Eds.), *Women and mathematics: Balancing the equation* (pp. 95-121). Hillsdale, NJ: Lawrence Erlbaum Associates.

Eccles, J. (1986). *Gender roles and Women's Achievements*, San Francisco: AERA.

Fenema, E., & Sherman, J. A. (1978). Sex-related differences in mathematics achievement and other factors: A further study. *Journal for Research in Mathematics Education, 9,* 189-203.

Granstam, I. (1985). *Measures to arouse girl's interest in technology.* Oxford: Gasat.

Greenfield, L. B., Holloway, E. L., & Remus, L. (1982). Women students in engineering. Are they so different than men? *Journal of college students personnel, 6*(23), 508-511.

Hanna, G. (in press). Cross cultural differences in mathematics education. To appear in: *International Journal of Educational Research.*

Hoeven, A. van der, & Jungbluth, P. (1988). Leerkrachtenverwachtingen, streefniveaus en het advies voor voortgezet onderwijs. *School en Begeleiding, 5*(20), 11-13.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139-155.

Jagacinsky, C. M., & Lebold, W. (1981). A comparison of men and women undergraduate and professional engineers. *Engineering Education, 72*(3), 213-220.

Jörg, T., & Man in 't Veld, M. (1987). *Populariteit van natuurkunde bij meisjes in het AVO,* Vakgroep Natuurkundedidaktiek Utrecht.

Kodde, D. A., & Theunissen, M. A. M. (1985). *De schoolloopbanen van snelle VWO-leerlingen, sexe-ongelijkheid en keuze van hoger onderwijs.* Economisch Instituut Nijmegen, Instituut voor Onderwijskunde, Ministerie van O&W. Den Haag: Staatsuitgeverij.

Kremers, E. J. J. (1990). *Overzicht van leerresultaten aan het einde van de eerste fase Voortgezet Onderwijs.* Arnhem: CITO.

Kristensen, D., & Jenneskens, A. (1991). *Emancipatiebeleid op schoolniveau.* Lisse: Swets & Zeitlinger.

Kuyper, H., & Meulenbeld, J. (1989). *De invloed van het gedrag van docenten op de prestaties in, keuze van en attitudes ten opzichte van wiskunde door meisjes in het AVO/VWO.* Groningen: RION.

Kuyper, H., & Werf, M. van der, (1987). *De invloed van het gedrag van docenten op wiskundegedrag door meisjes in het AVO, VWO.* Groningen: RION.

Kuyper, H. (1990). Sekseverschillen in het toekomstperspectief van MAVO-, HAVO- en VWO-leerlingen. *Jeugd en Samenleving, 20,* 540-546.

Lantz, A. E., & Smith, G. P. (1981). Factors influencing the choice of nonrequired mathematics courses. *Journal of Educational Psychology, 73*, 825-837.

Linn, M. C., & Hyde, J. S. (1989). Gender, mathematics and science. *Educational Researcher, 8*, 17.

Linn, M. C., & Peterson, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*, 1479-1498.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford: Stanford University Press.

Mullis, I. V. S., & Jenkins, L. B. (1988). *The science report card: Elements of risk and recorvery*. Princeton, NJ: Educational Testing Service.

Murphy, R. J. L. (1978). Sex differences in Examination Performance: do these reflect diffences in ability of sex-role-stereotypes? *Educational Review, 30*, 259-263.

Niphuis-Nell, M. (1992). *De emancipatie van meisjes en jonge vrouwen*, cahiernr. 88. Den Haag: VUGA

Oakes, J. (1990). Opportunities, Achievement, and Choice: Women and Minority Students in Science and Mathematics. *Review of Research in Education, 16*, 153-222.

Thomas, G. (1984). *Determinants and motivations underlying the college major choice of race and sex groups*. Baltimore: Johns Hopkins University, Center for the Social Organiszation of Schools.

Vonderen, N. van, & Raaf, J. de (1981). *De TH als studiekeuzemogelijkheid voor meisjes*, Eindhoven: Hogeschool Eindhoven.

Walberg, H. J. (1991). Improving School Science in Advanced and Developing Countries. *Review of Educational Research, 61*(1), 25-69.

Ware, N. C., & Lee, V. (1985). *Predictors of science major choice in a national sample of male and female college students*. Unpublished manuscript, Radcliffe college, Cambridge, MA.

Ware, N. C., Steckler, N. A., & Leserman, J. (1985). Undergraduate women. Who choose a science major? *Journal of Higher Education, 56*(1), 73-84.

Ware, N., & Dill, D. (1986). *Persistence in science among mathematically-able male and female college students with pre-college plans for a scientific major*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Weiner, B., Frieze, J. H., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R.M. (1971). *Perceiving the causes of success and failure*. Morristown, NJ: General Learning Press.

Willett, J. B. (1988/89). Measurement of change. *Review of Research in Education, 15*, 345-422.

Yee, D. K., Jacobs, J., & Goldsmith, R. (1986). *Sex equity in the home: Parents' influence on their children's attitudes about math*. Paper presented at the annual meeting of the American Educational Research Association, San Fransisco.

Zimmerer, L. K., & Bennett, S. M. (1987). *Gender differences on the California statewide assessment of attitudes and achievement in science*. Paper presented at the annual meeting of the American Educational Research Association, Washington DC.

Manuscript aanvaard 30-08-1993.

Auteur

H. Dekkers is senior-onderzoeker en plaatsvervangend hoofd van de afdeling Onderwijskansen (OKA) van het ITS.

Adres: Instituut voor Toegepaste Sociale Wetenschappen, Postbus 9048, 6500 KJ Nijmegen.

Abstract

Determinants of gender-related subject choices in secondary education

H. Dekkers. Pedagogische Studiën, 1993, 70, 445-459.

This article reports the results of a longitudinal study on the development of gender differences in subject choice, i.c. choice of mathematics (and science). Both boys and girls are followed between their 12th and 15th year of age, concerning (the development in) their attitudes, attributions, achievement and choice behaviour. A model is developped that shows the relation between changes in certain predictors (attitudes) and changes in subject choice, accounted for gender- and schooltype-differences.

Kroniek

Het AERA-Congres 1993, Atlanta

Introductie (N. Verloop, Rijksuniversiteit Leiden)

Het jaarlijkse AERA-congres werd dit jaar gehouden in Atlanta. Ook dit jaar weer waren, in deze bij uitstek Amerikaanse stad van Coca Cola en CNN, de niet-Amerikanen ruimschoots vertegenwoordigd, niet alleen als congresbezoekers maar ook als paperlezers.

Het thema van dit jaar was "The art and science of educational research and practice". Als selectiemiddel voor de inhoud van hetgeen men te verwachten heeft bij een AERA-congres heeft zo'n congresthema weinig te betekenen. De AERA heeft al een lange traditie van relatieve autonomie van de verschillende divisies. Binnen deze divisies wordt over lopend of recent afgesloten onderzoek gerapporteerd en dit laat zich uiteraard niet sturen op basis van een thema dat door de congresleiding wordt vastgesteld. Op een andere wijze is echter de keuze voor een dergelijk thema wel van belang, namelijk als een treffende illustratie van de enorme diversiteit in de duizenden bijdragen die te beluisteren waren. Het thema geeft immers aan dat zowel 'research' als 'practice' aan de orde waren en dat deze bovendien nog konden worden opgevat als zowel 'art' als 'science'. De 'program chair', Donmoyer, maakte van deze nood een deugd door te stellen dat juist in deze enorme diversiteit de kracht van deze conferentie gezocht moest worden en dat althans deze diversiteit als iets consistents in de AERA-congressen kon worden gezien ("A celebration of diversity"). Een onmiskenbaar voordeel van een AERA-congres is inderdaad dat elke congresganger in dit enorme aanbod wel een aantal waardevolle en bij zijn of haar directe interesse aansluitende zaken kan vinden. Voorwaarde is grondige bestudering vooraf van het lijvige programmaboek (waarin alleen al 400 pagina's nodig zijn om de titels van de bijdragen op te sommen).

Wat de vormgeving betreft, het experimenteren met nieuwe formats die afwijken van de traditionele paper-, symposium en poster-presentaties gaat nog steeds door. De als 'experimental' aangeduide sessies hadden alle tot doel een meer intensieve interactie tussen de 'toehoorders' te bewerkstelligen, hetgeen uiteraard slechts aangenaam is als deze toehoorders iets substantieels te melden hebben.

In de hier volgende rapportage is, net zoals enkele jaren geleden al eens gebeurd 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 geweest is. Aan de auteurs is gevraagd vooral in te gaan op de algemene trends die binnen de divisies zichtbaar waren. Van de 11 divisies komen er 5 niet aan bod. Dit betreft de wat minder omvangrijke of nogal specifiek Amerikaans ingevulde divisies. De 6 besproken divisies bestrijken het overgrote deel van hetgeen in het Nederlandse taalgebied onder 'onderwijskunde' of 'onderwijswetenschappen' wordt verstaan.

Divisie B: Curriculum studies (J. J. H. van den Akker, Universiteit Twente)

Het meest opvallende kenmerk van het programma van de Curriculum-divisie betrof dit jaar niet de inhoud maar de vorm. De variatie aan onderwerpen was zo mogelijk nog bonter dan anders; duidelijke inhoudelijke zwaartepunten heb ik niet kunnen ontdekken. Nadrukkelijk afwijkend was de alternatieve wijze waarop het aanbod gestructureerd was. Bij de (39) symposia (sets van gecoördineerde presentaties) werd dikwijls geëxperimenteerd met manieren om het zaalpubliek nadrukkelijk bij de discussie te betrekken. De gebruikelijke papersessies (waarbij een stuk of vijf losse papers, met vaak twijfelachtige samenhang, onder één noemer worden samengebracht, waarna een discussiant de ondankbare rol heeft iets zinnigs over het geheel te zeggen) waren dit maal achterwege gelaten. In plaats daarvan was (in samenwerking met Division G: Social Context of Education) een grote eendaagse

Poster Fair georganiseerd waarin de auteurs van alle 85 curriculumpapers zich afzonderlijk via hun posters konden manifesteren en met geïnteresseerden in discussie treden. Naar mijn indruk is die formule goed bevallen. Vooral de interactie tussen de deelnemers - en daar is zo'n conferente toch vooral voor bedoeld - was aanzienlijk groter dan bij de meer traditionele aanpak. Mij persoonlijk viel op dat de wat opener werkwijze leidde tot meer contacten en uitwisseling van papers en ervaringen met de vele aanwezige onderzoekers uit andere, niet-Amerikaanse, werelddelen. De drempel tot communiceren ligt bij postersessies blijkbaar wat lager dan bij plenaire discussies, waarin de Amerikaanse deelnemers veelal domineren.

In hoeverre men zich inhoudelijk aangesproken voelt door het aanbod, hangt vanzelfsprekend sterk af van de persoonlijke interesse. Ook dit jaar was kenmerkend voor veel van de Amerikaanse curriculumpresentaties (men denke met name aan het werk van auteurs als Apple, Giroux e.d.) de filosofische en ideologische inslag en de kritisch-beschouwende benadering, die de papers veeleer op essays dan op onderzoeksrapportages doen lijken. De relatie met de meer praktische problemen rondom curriculumontwikkeling komt dikwijls niet of zeer bescheiden aan bod. Dat wijkt enigszins af van de hoofdstroom in het Nederlandse curriculumonderzoek, waar empirisch onderzoek een grotere rol speelt en het laatste decennium veel aandacht is besteed aan meer pragmatische vraagstukken rond curriculumontwerp, -evaluatie en -implementatie. Duidelijke sporen daarvan kan men aantreffen in het recente 'Handboek Curriculum: modellen, theorieën, technologieën, (onder redactie van W. J. Nijhof et al., uitgegeven door Swets & Zeitlinger).

Die geringe wisselwerking tussen curriculumtheorie en onderwijspraktijk in veel van de Amerikaanse curriculumpresentaties op de AERA is voor een deel te wijten (zo luidt althans de verklaring van een Amerikaanse collega) aan het feit dat veel van de auteurs docent zijn aan *graduate schools* waarin met name de Ph.D. (of Ed.D) studenten primair opgeleid worden voor universitaire posities en niet voor banen op het terrein van curriculumontwikkeling of iets dergelijks in de *real world*. Een dergelijke oriëntatie herbergt het risico van een zichzelf reproducerende, ietwat inteeltachtige wetenschappelijke discussie.

Een andere consequentie van de geringe hoeveelheid op de praktijk gericht onderzoek is dat auteurs van handboeken op het terrein van curriculumontwikkeling hun voorstellen nauwelijks kunnen baseren op resultaten van empirisch onderzoek naar ontwikkelstrategieën en effecten daarvan. Een succesvol auteur op dit terrein vertelde me dat zijn uitgever (uit commerciële motieven) grote druk uitoefende om met een nieuwe, gereviseerde versie op de markt te komen, maar dat hij daar zelf grote reserves bij had. Immers, op welke empirisch gefundeerde uitspraken in de literatuur zou hij nieuwe voorstellen moeten baseren?

In dit verband is het interessant terug te grijpen op een symposium dat vorig jaar (AERA-conferentie, 1992, San Francisco) belegd werd ter gelegenheid van het verschijnen van het 'Handbook of research on curriculum' (met Philip Jackson als editor, uitgegeven door Macmillan). Hoewel dat alleen al door zijn omvang (1088 pagina's) imponerende boek een zeer waardevol overzicht bevat van het curriculaire kennisbestand, luidde een door meerdere discussianten geuit kritiekpunt dat de betoogtrant in veel hoofdstukken wel erg afstandelijk en beschouwend was en dat het boek wel erg weinig praktisch houvast bood. En dat terwijl Jackson in zijn inleidende hoofdstuk (Conceptions of curriculum and curriculum specialists - boeiend voor de fijnproevers) schrijft over de toekomst van "curriculum as a field endeavor whose central professional task is that of giving advice to practitioners as they try to improve the curriculum" (p. 34). Vanuit een dergelijk gezichtspunt is naar het oordeel van sommigen in het boek te weinig aandacht besteed aan praktisch relevante onderwerpen als: de rol van uiteenlopende deelnemers aan curriculumontwikkeling, de positie en het functioneren van leraren (het bestuurlijke perspectief zou overheersen), curriculaire problemen in schoolvakken, en concrete procedures voor curriculumontwikkeling. Jacksons reactie op de kritiek tijdens het symposium was enigszins relativerend: er was geen praktisch *how-to-do* boek nagestreefd, het curriculumdomein is nu eenmaal vol onzekerheden, en men dient zich te realiseren dat dit pas het eerste (maar waarschijnlijk ook het laatste!) curriculum-

handboek was dat deze eeuw verschijnt. Althans in Amerika, zullen wij er zelf ter verdere relativering maar aan toevoegen, want naast het bovengenoemde, recente Nederlandse handboek kunnen we bijvoorbeeld ook wijzen op het in 1983 verschenen kloeke Duitse 'Handbuch der Curriculumforschung' (onder redactie van Hameyer et al., uitgegeven door Beltz).

Het grote voordeel van zo'n brede en gevarieerde conferentie als die van de AERA is natuurlijk dat je als curriculumonderzoeker veel interessants kunt meemaken in andere divisies, waar vaak vele raakvlakken met het curriculumdomein zijn, zoals in de divisies A: Administration (denk aan inhoudelijk onderwijsbeleid), divisie H: School evaluation and program development (de overlap is evident) en in divisie K: Teaching and teacher education (denk aan de wisselwerking tussen curriculumontwikkeling en professionele ontwikkeling van leraren). En curriculumonderzoekers met een sterke interesse in ontwerpvraagstukken kunnen soms het nodige van hun gading vinden in divisie C (Learning and instruction), vooral op het vlak van *instructional design*, hoewel daar soms wel weer erg luchthartig wordt omgesprongen met vraagstukken rondom enerzijds doelbepaling en anderzijds implementatie van vernieuwingen. Vanuit een curriculair perspectief klinken de soms stellige beweringen van instructie-onderzoekers over de doeltreffendheid van hun oplossingen soms enigszins naïef.

Enkele afzonderlijke thema's die mij dwars door het programma heen zijn opgevallen, betreffen:

* *(Interactieve) video:* meer dan in vorige jaren was er een sterke aandacht voor het benutten van video in de instructie van leerlingen, maar ook voor de (na)scholing van leraren (in opleiding). Opvallend was ook de frequentie waarmee videofragmenten in de presentaties zelf benut werden (en dat op een conferentie waarin het gebruik van de overheadprojector in recente jaren nog tamelijk gering was).
* *Alternative assessment:* vanuit een veelheid aan perspectieven (op terreinen als: curriculum, instructie, evaluatie, toetsing, innovatie, technologie) groeit de belangstelling voor andere dan 'traditionele' toetsvormen. Allerlei verwante termen zijn in omloop: *performance/ authentic/ portfolio assessment*. Een basismotief lijkt de overtuiging dat 'alternatieve' vormen van het bepalen van leervorderingen kunnen bijdragen aan onderwijsverbetering. Ondanks de variatie aan begrippen en invalshoeken lijkt er sprake van convergerende benaderingen die de komende jaren volop in de schijnwerpers zullen staan.
* Veel aandacht voor zogenaamde *collaborative projects:* vormen van samenwerking tussen universitaire onderzoekers en scholen/leraren, waarin gezamenlijk gestreefd wordt naar praktijkverbetering. Sommige cynici spreken uit dat de motieven voor dergelijk actie-onderzoek - om maar eens een al weer wat gedateerde term te gebruiken - vooral samenhangen met verminderde onderzoeksbudgetten die dwingen tot meer kleinschalig opereren. Mits die projecten, door een uitgekiende en methodologisch zorgvuldige aanpak, ook leiden tot wetenschappelijke kennisgroei, lijkt me dat echter een goed voorbeeld van hoe men van de nood een deugd kan maken.

Enigszins verwant aan deze laatste trend (koppeling van onderzoek en praktijkverbetering) was een vanuit de UT (Plomp en Van den Akker) in samenwerking met Walker (Stanford University) georganiseerd symposium over 'Development research in curriculum', waarin voorstellen over ontwerpgericht ofwel ontwikkelingsonderzoek op het curriculumdomein gepresenteerd en bediscussieerd werden. Vooral de wijze waarop het ontwerpen en evalueren van kwaliteitsrijke curriculumprodukten gepaard kan gaan met het genereren en beproeven van curriculumstrategische kennis stond daarbij centraal. Ook hier lijkt sprake van een benadering die de komende jaren veel (hernieuwde) belangstelling zal ondervinden vanuit zowel wetenschappelijke als beleidsmatige en praktische kringen. Enkele kernthema's in dit verband zijn: systematisering van ontwerpbenaderingen, intensivering en benutting van formatieve evaluatie, verheldering van functies van curriculumprodukten binnen dikwijls veelomvattende innovatieprogramma's, efficiency-overwegingen bij ontwikkelstrate-

gieën en gerichte toepassing van informatietechnologische hulpmiddelen.

Voor wie naar aanleiding van deze schets de verzuchting mocht slaken dat het curriculumdomein wel erg breed en vaag is, besluit ik met een ander citaat van Jackson (o.c., p. 37): "The boundaries of the field are diffuse, so much so that one may wonder sometimes whether it has any boundaries at all. To some, that condition is troublesome; to others, it is exhilarating; to all, it can become confusing at times".

Divisie C: Learning and instruction (S. Dijkstra, m.m.v. T. de Jong en J. J. G. van Merriënboer, Universiteit Twente)

Algemeen

Tijdens de AERA annual meeting viel het opnieuw op dat in vergelijking met onze grote buurlanden, de Nederlandse onderwijskundigen hun wetenschappelijke produktie met name aanbieden aan Amerikaanse collega's. Dit geldt eveneens als een vergelijking met Japan gemaakt wordt. Het aantal Nederlandse deelnemers was overigens zeker het dubbele van het aantal dat een presentatie verzorgde. Het is duidelijk dat de Nederlanders vrij veel gewicht hechten aan de onderwijskundige theorieën die in de Verenigde Staten van Amerika ontwikkeld worden. In Tabel 1 wordt een overzicht gegeven van de aantallen personen uit enkele ons omringende landen en Japan die op de een of andere wijze - symposium, lezing, poster, e.d. - een bijdrage op de AERA verzorgd hebben. Ongeveer 95% van de bijdragen worden verzorgd door personen, die uit de Verenigde Staten en Canada afkomstig zijn. Uit die twee landen betrof dat dit jaar in totaal 3896 personen, terwijl het totale aantal bijdragen 4108 was. De gegevens uit Tabel 1 dienen tegen deze achtergrond geïnterpreteerd te worden.

Tabel 1
Aantal personen uit enkele Europese landen en Japan, dat in 1991, '92 en '93 op de AERA annual meeting een presentatie verzorgd heeft

	1991	1992	1993
Engeland	37	42	24
Duitsland	12	12	10
Japan	10	3	1
Zweden	12	6	5
Noorwegen	5	7	9
Nederland	33	35	38

In vergelijking met de jaren zestig en zeventig is er de laatste tien jaar sprake van een volledige ommezwaai in de bestudering van leren en instructie. In de eerstgenoemde periode lag de interesse in het volgen van het probleemoplossen en leren met name op het effect van 'feedback' en het proces van informatieverwerking. De problemen die werden bestudeerd waren doorgaans goed-gestructureerd, de interpretatie van het proces vond plaats binnen de psychologie van de informatieverwerking, de instructie werd vorm gegeven door ontwerptheorieën, waarin 'systeem controle' een belangrijke rol speelde. Thans ligt de nadruk op de bestudering van de constructie van kennis door de leerling en op vormen van 'vooruitkoppeling' in de instructieve programma's. In de programma's krijgen de leerlingen mogelijkheden aangeboden om in verschillende richtingen verder te gaan, afhankelijk van verschillende doelen of niveaus van interesse in het gebied. De interpretatie van het verloop van het leerproces vindt plaats binnen constructivistische ideeën over verwerving van kennis en vaardigheden en binnen theorieën over leerlingcontrole en zelfsturing.

Hoewel de zelfstandige ontwikkelingen in de cognitieve psychologie de ommezwaai beïnvloed hebben, is deze zeker mede geïnspireerd op de ontwikkelingen in de informatietechnologie. Het beschikbaar komen van programma's als 'Hypertext', de multimedia aanpak, de toename in rekensnelheid van de computer, gecombineerd met de miniaturisering van het apparaat, waardoor deze op grote schaal bruikbaar werd in het onderwijs, hebben het mogelijk gemaakt de aandacht veel meer te richten op het begin van het onderwijsleerproces en op het exploreren van het instructieve materiaal. Tijdens sommige AERA poster-

symposia werden multimediale programma's gedemonstreerd. Hoewel in het onderwijs zeker gedurende de laatste 100 jaar van verschillende media gebruik gemaakt wordt houdt de huidige multimedia 'revolutie' in, dat alle vormen van informatie, zowel visueel (stilstaand en bewegend), als grafisch en ook auditief, digitaal opgeslagen kunnen worden in één format en daarmee bruikbaar zijn voor gecombineerde computergestuurde presentatie. De programma's die getoond werden zijn afkomstig van twee onderzoek- en ontwikkelcentra, het Peabody College van Vanderbilt University, Nashville, TN en het Institute for the Learning Sciences van Northwestern University, Evanston, IL. Deze instituten hebben ruime subsidies verkregen om de multimediale produkten te realiseren. De inhoud ervan betreft vooral geografie en geschiedenis, maar ook het rekenonderwijs. Voorbeelden zijn het programma 'Road Trip', waarin de leerling op het computerscherm een wegenkaart van een Amerikaanse staat ziet. Hij of zij kan met behulp van de muis een weg afrijden en in een bepaalde plaats stoppen. Op de video kan de leerling dan één of meer activiteiten, die specifiek voor die plaats zijn, bekijken en de relevante informatie opnemen. De leerling is geheel vrij in het kiezen van de weg, het leren vindt explorerend plaats. Een ander programma 'Advise the President' heeft de recente Amerikaanse geschiedenis tot onderwerp. De leerling wordt interactief betrokken in de crisis in de betrekkingen tussen de Verenigde Staten van Amerika en Iran. De leerling kan, als er in het diplomatieke proces een volgende beslissing genomen moet worden, oud-president Carter adviseren welke beslissing te nemen. De president antwoordt waarom de beslissing wel of niet genomen kan worden en wat de gevolgen van een beslissing kunnen zijn.

Het kernbegrip bij bovenstaande multimediale aanpak is 'anchored instruction'. Een situatie wordt als ankerpunt genomen van waaruit het exploratieve leerproces plaats vindt en de leerling zijn ontdekkingen kan doen. In de discussies tijdens en na afloop van de poster demonstraties viel het op dat er behoefte was aan een duidelijke doelformulering over welke kennis en vaardigheden de leerlingen zouden moeten verwerven. Verder werd er sterke behoefte gevoeld aan het beschikbaar komen van evaluatie-gegevens: Wat is er daadwerkelijk geleerd en hoe zou dat geëvalueerd moeten worden. Ook zou standaardisatie in programmatuur en apparatuur wenselijk zijn voordat de uitgevers daadwerkelijk geïnteresseerd zijn om het materiaal aan het onderwijs aan te bieden.

Het is vrijwel niet mogelijk om een compleet overzicht te geven van alle ontwikkelingen binnen de Divisie Leren en instructie. Enerzijds is het lastig om een goed indelingscriterium te vinden. Een indeling naar sector voldoet niet om de ontwikkeling in theorievorming adequaat te schetsen. Daar komt nog bij dat leren en instructie voor specifieke vakgebieden (Education in the professions) in een andere divisie is ondergebracht en verder ook interessante ontwikkelingen geschetst worden in divisie K: Teaching and teacher education. De ontwikkelingen in theorievorming en instructievormgeving voor het leren van volwassenen komen op de AERA slechts in beperkte mate aan de orde, omdat de bedrijfsopleiders in dezelfde week waarin de AERA plaats vindt hun eigen 'annual meeting' hebben. Indelingen naar vakgebied zijn ook niet geheel bevredigend, enerzijds omdat de theorievorming dikwijls sterk specialistisch wordt, anderzijds omdat het lastig is het aantal vakgebieden precies af te bakenen. Vandaar dat volstaan wordt met een korte beschrijving van een aantal symposia, waarin de nieuwe theoretische en praktische trends zich helder aftekenen.

Constructivistische leeromgevingen: practica, groepswerk en 'microworlds'

Zoals hierboven reeds is vermeld was er op deze AERA voortgaande aandacht voor de constructie van kennis als fundamenteel proces van kennisontwikkeling door de leerling en voor het ontwerpen van constructivistische leeromgevingen. Echter, de eerder sterk aanwezige verbinding tussen constructivisme en situationisme wordt steeds meer losgelaten. Constructivisme wordt dan een actieve, zelf gerichte vorm van leren, waarbij het verkrijgen van 'objectieve' kennis niet wordt uitgesloten. Zelfs situationisme als begrip op zich krijgt een wat minder uitgesproken betekenis in die zin dat het nu meer en meer gaat betekenen dat het kennisacquisitieproces in een reële (of op de realiteit gelijkende) situatie plaats vindt. Met

name Collins benadrukte een aantal keren dat je moet leren voor transfer en dat gesitueerde kennis 'geglobaliseerd' moet worden. Wel vindt hij dat er van gesitueerde kennis naar abstracte kennis moet worden gegaan en niet andersom, zoals in traditionele curricula vaak het geval is.

De boven omschreven trend heeft geleid tot een groot aantal onderzoeken en projecten waarin de lerende zich als een 'wetenschappelijk' onderzoeker moet gedragen. De student of leerling moet hypothesen opstellen, data verzamelen, grafieken interpreteren, discussiëren met 'collega's', conclusies trekken en rapporten schrijven. Vaak wordt dit proces ondersteund door technologieën als computersimulatieprogramma's maar ook door allerlei proefopstellingen. Hieronder volgen beschrijvingen van symposia en lezingen die min of meer illustratief zijn.

Het symposium 'Issues of cooperative learning and learner control in computer-based and other media environments' valt binnen bovengenoemd kader. De relatie tussen coöperatief leren en leerlingcontrole bleef daarbij overigens achterwege. Een aannemelijke hypothese zou kunnen zijn dat leerlingcontrole meer effectief is wanneer in twee- of drietallen gewerkt wordt. Er werd echter een viertal studies gepresenteerd (Ross, Schick, Weiser, Klein), die of coöperatief leren of leerlingcontrole bestudeerden. Uit de inhoud van het symposium 'Instructional applications of hypermedia: Profiles of interactive student learning' werd eveneens duidelijk dat het onderzoek naar hypertekst en hypermedia meer en meer gebaseerd is op constructivistische noties, zoals de Cognitive Flexibility Theory.

In het symposium 'Is programming obsolete in the classroom?' met bijdragen van DiSessa, Resnick en Smith werd duidelijk dat programmeren in engere zin (Basic, Logo etc.) geen hot topic meer op de AERA is. Maar de presentatoren demonstreerden of beschreven met veel enthousiasme enige nieuwe paradigma's voor programmeren (naast BOXER van DiSessa ook programmeren door directe manipulatie, bouwstenen voor animatie, intelligente Lego-blokjes etc.). Deze systemen hebben meer met 'Instruction by Design' (IbyD) dan met traditioneel programmeren te maken; de beschreven paradigma's passen goed in een constructivistische onderwijsopvatting en het valt derhalve te verwachten dat zij zich binnenkort in een grotere populariteit mogen verheugen dan heden het geval is. De klassieke idee dat het leren probleemoplossen op zich gestimuleerd wordt door 'leren programmeren' is niet losgelaten.

In een van de symposia waar het gebruik van een constructivistische leeromgeving tot uitdrukking kwam werd een onderwijsprogramma voor het onderwijzen van seizoenen en de stand van de zon en aarde ten opzichte van elkaar behandeld. In dit programma spelen (zelfgemaakte) video's maar ook eenvoudige voorwerpen als een grote (aard)bol, een stokje dat hierop geplaatst kan worden en een lamp een rol, die tezamen als een soort 'microworld' dienst doen. Essentieel is dat leerlingen zelf hypothesen opstellen, data verzamelen en die met elkaar bediscussiëren (beweerd werd dat in een gewone klas slechts 5% van de tijd aan discussie besteed wordt). De docent speelt bij de discussie slechts de rol van facilitator, alleen wanneer de discussie dreigt vast te lopen geeft hij een 'hint' om verder te gaan. In het getoonde voorbeeld plaatsen de leerlingen het stokje boven op de Noordpool. Toen de discussie dreigde te verzanden gaf de docent slechts een vage hint dat dit een speciaal geval was en dat ze ook andere plaatsen konden bekijken. Videobanden van dit soort klassen worden vanuit een aantal perspectieven geanalyseerd: de 'discourse', de 'gender' en de kennisontwikkeling. Schauble van de Universiteit van Wisconsin gaf haar visie op het laatste en zij vond het opvallend dat de modelontwikkeling van de leerlingen op de band niet plotseling en bewust was maar eerder, wat zij noemde, 'slippery' was.

Ook bij abstractere vakgebieden als wiskunde wordt gebruik gemaakt van 'microworlds'. De groep rond Scheffe aan de universiteit van Georgia presenteerde een aantal microworlds voor wiskunde. 'Sticks' is bijvoorbeeld een microworld voor leren rekenen waarin kinderen stokken kunnen breken en combineren waarmee getracht wordt hun al spelenderwijs een idee van eenheden te geven. Ook hier speelt de docent slechts een faciliterende rol. Overigens was het opvallend dat bij de getoonde video-opnames er altijd één docent op twee kinderen aanwezig was.

Een bijzonder interessant symposium was

georganiseerd door de Boston University. Dit symposium viel op zowel door de inhoud als door de opzet. Het ging hier om een 'science curriculum' waarin leerlingen met behulp van fractals (beschrijven van gestructureerde groeipatronen vanuit random gebeurtenissen) processen in verschillende domeinen konden analyseren en modelleren. In het symposium werd specifiek gekeken naar electrochemische verschijnselen, waarbij de groei van het neerslaan van koper uit een kopersulfaatoplossing bij een kathode werd getoond. Leerlingen maken de fysieke proefopstelling en kunnen vervolgens het resultaat van de proef (het patroon van de neerslag) op video vastleggen. Dit videobeeld kunnen zij in de computer inlezen waarna het patroon kan worden geanalyseerd en er kunnen simulaties mee worden uitgevoerd. Ook worden simulaties gebruikt van andere processen (bijvoorbeeld groei van bacteriekoloniën), en kunnen leerlingen zelf parameters in de simulatie veranderen. In het programma ligt een sterke nadruk op het doen van voorspellingen voor er met de proefopstelling en de simulatie wordt gewerkt. De groep onderzoekers noemt dit 'learning through science'. Een observatie van het traditioneel curriculum met de nieuwe opzet liet zien dat het percentage 'lecturing' zakte van 91% naar 10%. Daarentegen namen coaching en scaffolding activiteiten van de docent drastisch toe. Gemeten met 'concept maps' bleek de experimentele groep een kennisstructuur te vertonen die dichter bij die van experts lag dan de traditionele groep. Het aardige van het symposium was de compleetheid. Er was een korte introductie tot het project, de natuurkundige proef werd life uitgevoerd, op video opgenomen en ingelezen. De deelnemers zagen video opnames van kinderen die aan het project deelnamen. Een onderzoekster vertelde hoe zij een en ander empirisch had onderzocht en een groep docenten vertelde hun ervaringen. Ten slotte gaf Collins als discussiant zijn commentaar.

Naast dit soort omvattende projecten waarin een (deel)curriculum wordt aangepakt en de onderzoekscomponent een minder prominente rol speelt was er een groot aantal presentaties waarin een van de aspecten van constructivistisch leren gedetailleerd bekeken werd. Vragen die aan de orde komen zijn: welke soort toetsen moeten gebruikt worden bij constructivistisch leren (Collins), hoe gaan leerlingen om met onvoldoende data (Lewis), of met data die in tegenspraak met de huidige theorie zijn (Chinn & Brewer) en hoe genereren leerlingen (meerdere) hypothesen en welk effect heeft dit op wat ze leren van een simulatieomgeving (Quinn & Alessi). Het is duidelijk dat het onderzoek naar exploratief leren en alle processen die daarbij horen aan het toenemen is.

Expertiseniveau en kennisrepresentatie

Het onderzoek naar kennisrepresentaties op verschillende expertiseniveaus blijft een interessant en frequent thema op de AERA. Met name op het gebied van de geneeskunde is een aantal onderzoeksgroepen zeer actief, waaronder die rond Schmidt en Boshuizen te Maastricht. Ook dit jaar was er weer een aantal presentaties uit deze groep onder andere in een symposium onder de titel 'Studies in professional expertise'. Tijdens dit symposium vond een interessante methodologische discussie plaats. Eén van de fenomenen die met name bij geneeskunde gevonden zijn is het zogenaamde 'intermediate-effect'. Om kennismodellen te meten worden personen 'cases' aangeboden en later wordt hun gevraagd de case te reconstrueren. In tegenstelling tot wat verwacht zou kunnen worden blijken 'intermediates' betere reconstructies te geven dan experts (en beginners). Dit wordt verklaard vanuit een sterkere gecompileerdheid (of ge-encapsuleerdheid) van kennis bij experts. In een van de presentaties bleek dat het intermediate-effect niet altijd optrad en dit werd toegeschreven aan het soort cases dat gebruikt werd (a-typische cases). Tijdens de discussie bleek echter dat experts bij reconstructie vaak labels gebruiken die staan voor een grotere hoeveelheid gegevens en dat bij het scoren van de reconstructies deze labels normaal gesproken slechts één keer meetellen. Intermediates bezitten deze 'chunks' nog niet en reconstrueren daarom in alle volheid. Het intermediate effect zou dus wel eens een artefact van de scoringswijze kunnen zijn. Feltovich wees in zijn commentaar nog eens op de flexibiliteit die kenmerkend is voor experts, indien daar specifiek om gevraagd zou worden, zouden deze dus wel een volledige reconstructie kunnen geven.

Leertheorie en het ontwerpen van instructie

De relatie tussen leertheorie en instructieontwerp bleek nogal problematisch tijdens het symposium 'A dialogue on the relationship of educational learning theory to instructional theory'. Een belangrijke oorzaak hiervoor leek te zijn dat de vier presentatoren (Breuer, Tennyson, Blumenfeld, Driscoll) ieder iets anders verstonden onder 'educational learning theory', variërend van cognitieve leerpsychologie, via sociaal-constructivistisch leren, tot semiotische opvattingen van het onderwijsleerproces: dat maakt een diepgravende discussie over de relatie tot de instructietheorie erg moeilijk.

Min of meer verwant aan het vorige symposium was het symposium 'Are we ready for a cognitive revolution in education? Is there going to be one? What will it be like when and if it happens?' Alle presentatoren (o.a. Beck, Palincsar, Bransford, Pressly, Swanson) waren het er over eens dat de cognitieve revolutie in de psychologie weliswaar heeft plaats gevonden, maar dat deze nog nauwelijks enige uitwerking heeft gehad naar de praktijk van het onderwijs - terwijl dat wel zou moeten. Het symposium bestond voor het grootste deel uit een vraag-en-antwoord-spel met het publiek. Een aardige opzet, maar de felheid waarmee de Amerikanen steeds weer de discussie voeren over (de noodzaak van) 'school reform' doet voor een Europeaan wat vreemd aan.

De ontwikkelingen in en de eigenschappen van het ontwerpen van instructie kwamen uitgebreid aan de orde in het symposium 'Instructional design as ill-structured problem solving: implications for theory and practice' (Van Merriënboer, Tennyson, Breuer, Gros & Dijkstra). Tennyson schetste de ontwikkelingen van het instructieontwerp vanaf de jaren zestig en kwam tot een indeling in vier generaties van instructieontwerp, waarbij hij ook op de automatisering van de ontwerpactiviteit inging. Van Merriënboer schetste de probleemruimte voor het instructieontwerp en concludeerde dat er slechts in weinig gevallen een algoritme gebruikt kan worden voor het ontwerpen van instructie, maar dat er veeleer sprake is van heuristische methoden. Deze zienswijze heeft implicaties voor de theorie en praktijk van het ontwerp, waaronder het formaat en de organisatie van ontwerpmodellen en de automatisering van gedeelten van het ontwerpproces. In de discussie ging Dijkstra in op de aard van de communicatie in instructieve omgevingen, op de representatie van de inhoud van het leermateriaal en op de vraag welke vorm van interactiviteit samen hangt met de typen door de leerling op te lossen problemen.

Bij de ledenbijeenkomst van de SIG Instructional Technology werd een presentatie verzorgd door Schott van de Universiteit van Tuebingen, met de titel 'The architecture of instructional theory: problems in building a theory of instruction'. In de lezing werd ruim aandacht besteed aan de relatie tussen leerprocessen en instructietheorie, maar minder aan de ontwikkeling van een technologie vanuit die theorie. Dit leidde tot een levendige discussie met het publiek.

Het symposium 'Current issues in military training: What directions for instructional design and technology?' bood een aantal lezingen van hoge kwaliteit. Centrale elementen betroffen empirische studies naar het instructieontwerpproces (Perez), automatisering van dit proces met behulp van computer-ondersteunde systemen (AIDA: Spector), en toepassingen van 'virtual reality' binnen militaire training (Thurman).

Besluit

Ook dit jaar had de divisie 'Leren en instructie' enkele gastsprekers uitgenodigd, waaronder Davydov van de Russische Academie van Onderwijswetenschappen. Zijn lezing had als onderwerp 'De invloed van Vygotsky op de theorie en praktijk van het onderwijs'. De lezing bood een breed overzicht en illustreert ook hoe in politiek lastige omstandigheden academisch werk gedaan werd. Confrey, Greeno, Cobb en McDermott gingen in een volgend symposium in op de invloed die Vygotsky op het onderzoek van het onderwijs in de Verenigde Staten van Amerika gehad heeft.

Tot slot zij opgemerkt dat in het bovenstaande de belangrijkste trends in het onderzoek en ontwikkelingswerk geschetst zijn. De ontwikkelingen binnen vakgebieden zoals lezen, rekenen, taalverwerving, wiskunde, fysica, e.d., die in vele postersessies en lezingen aan de orde kwamen zijn buiten beschouwing gebleven omdat het fysiek niet mogelijk was ze alle bij te wonen.

Divisie D: Measurement and research methodology (P. Sanders, Cito)

De indruk bestaat dat er de laatste jaren steeds meer overlap is tussen de onderwerpen van divisie D van de AERA en de NCME (National Council on Measurement in Education). Afgezien van een tiental lezingen over en toepassingen van hiërarchische lineaire modellen en een aantal lezingen over zeer specialistische onderwerpen, hadden de onderwerpen van divisie D voornamelijk betrekking op zaken uit de onderwijsmeetkunde. In de lezingen van de AERA eisten vier onderwerpen (ook) dit jaar de meeste aandacht op: adaptief toetsen, itembias, itemresponstheorie en 'performance assessment'. Gezien de toenemende capaciteit en kwaliteit van personal computers is de verwachting dat de computer steeds meer ingeschakeld zal worden bij het toetsen. Daarbij gaat de aandacht met name uit naar (varianten van) adaptief toetsen die het mogelijk maken om met veel minder items, dus veel sneller en goedkoper, betrouwbare schattingen van de vaardigheid van leerlingen te verkrijgen. Wat betreft het onderzoek naar itembias of dif (differential item functioning) heeft dit onderzoek zich de afgelopen jaren voornamelijk gericht op itembias bij dichotoom gescoorde vragen. Nu we echter ook in de U.S.A. een toename zien van andere vraagvormen, wordt ook onderzoek gedaan naar itembias bij polytoom gescoorde vragen. Veel lezingen op het gebied van de itemresponstheorie hadden zoals gewoonlijk het Rasch model tot onderwerp. Een interessante lezing was die van de Nijmeegse hoogleraar Roskam over de bijdrage van Europese psychometrici aan de verdere ontwikkeling van het Rasch model.

Een zeer groot aantal lezingen had betrekking op meetinstrumenten of toetsen waarbij de leerlingen zelf een respons moeten construeren. Deze 'constructed response' toetsen zijn een reactie op het gebruik van multiple choice toetsen waarbij de respons alleen maar bestaat uit het aanstrepen van het goede alternatief. Achter de term constructed response gaat een grote verscheidenheid aan toetsvormen schuil. Het betreft hier varianten van meerkeuzevragen (verwijder de redundante woorden uit een passage), het schrijven van een essay, het uitvoeren en beschrijven van een experiment en zeer complexe opdrachten zoals het bespelen van een instrument. Tot constructed response meetinstrumenten behoren ook de meetinstrumenten die aangeduid worden met de termen 'performance assessment', 'authentic assessment' en 'portfolio assessment'. De laatste drie termen benadrukken nog eens dat de opdrachten die de leerlingen moeten uitvoeren, ontleend zijn aan de praktijk (van de beroepsbeoefenaar). Deze meetinstrumenten zouden behalve een grotere (indruks)validiteit ook meer mogelijkheden bieden op het gebied van het meten van 'higher order thinking skills'. Voor wat betreft 'science' hebben met name Shavelson en zijn medewerkers de afgelopen jaren de nodige publikaties doen verschijnen. Of de voordelen van deze meetinstrumenten opwegen tegen de nadelen moet de komende jaren blijken. Zoals verwacht kon worden, blijken er vele problemen met betrekking tot de ontwikkeling, scoring en beoordeling van deze instrumenten te zijn. De drie genoemde aspecten impliceren namelijk de inzet van deskundige ontwikkelaars en beoordelaars. Die inzet heeft geleid tot een aanzienlijke toename van kosten voor de scholen.

Een van de leukste symposia was die over 'criterion-referenced measurement' (crm). Het symposium werd gehouden naar aanleiding van het feit dat het dertig jaar geleden was dat het beroemde artikel van Robert Glaser in de American Psychologist verscheen. Naast Glaser gaven ook andere bekende wetenschappers op dit gebied zoals Popham, Hambleton en Millman, een lezing. Hambleton gaf een overzicht van het gebied waaruit bleek dat ongeveer 75% van de publikaties tussen 1975 en 1985 verschenen was. Uit het feit dat er de laatste jaren nog maar zeer weinig gepubliceerd was, mocht men volgens Hambleton niet de conclusie trekken dat de ideeën die ten grondslag lagen aan of verwant waren met crm (beheersingsleren, inhoudelijke onderbouwing van zak-slaag beslissingen) geen ingang gevonden zouden hebben in het onderwijs. Met uitzondering van Millman zag elke spreker een toekomst voor 'criterion-references measurement: a celebrated and frequently elaborated concept needed for the 21st century'.

Divisie H: School evaluation and program development (B. P. M. Creemers, Rijksuniversiteit Groningen)

Divisie H: School evaluation and program development, is binnen de AERA de verzamelplaats voor wetenschappelijk opgeleide practici en misschien ook wel voor praktisch georiënteerde wetenschappers. Het is een van de grootste divisies met een veelheid aan onderwerpen. Er komen er elk jaar een paar bovendrijven, een aantal voor evaluatie maar vooral op het brede terrein van onderwijsverbetering.

Wat de evaluatie betreft kan dit leiden tot sessies waarin zakelijk verslag wordt gedaan van evaluatie van bepaalde programma's tot een oproep om ideeën te ontwikkelen ten aanzien van de zoveelste vernieuwingen en veranderingen binnen de evaluatiemethodologie.

Een voorbeeld van het eerste is de evaluatie van de onderwijsvernieuwingen in Chicago en de resultaten van de Magnet schools, maar vooral van de alternatieve benaderingen van onderwijsevaluatie waren vele voorbeelden in het programma aanwezig. Wat de onderwijsverbetering betreft, en soms de interpretatie van wat onderwijsvernieuwing zou moeten zijn, is er een voortdurende aandacht voor alle vormen van 'dropouts', 'students at risk' en 'special education', 'drug abuse in schools' en speciale programma's, zoals 'Success for All' van Slavin en 'Chapter 1' (zowel inhoud als evaluatie ervan). Op dit moment heeft in divisie H 'school restructuring' veel aandacht en verder een aantal praktische uitwerkingen van ideeën rondom onderwijseffectiviteit voor zover ze niet bij 'restructuring' zijn ondergebracht.

Op het terrein van onderwijseffectiviteit blijft opvallen de grote praktische georiënteerdheid van het Amerikaanse onderzoek. Dit betreft zowel onderzoek naar de schoolleider, de schoolorganisatie, als de hogere lagen van bestuur zoals het districtsniveau. Meer conceptuele problemen gerelateerd aan verdere theorievorming komen eigenlijk alleen maar vanuit het internationale gezelschap, zoals de bijdrage van Yin Cheong Cheng en Zuzovsky. In de bijdrage van Yin Cheong Cheng wordt een relatie gelegd tussen diverse theorieën op het terrein van onderwijsorganisatie en organisatie in het algemeen en onderwijseffectiviteit, te onderscheiden in efficiëncy en effecten. Op zich niet veel nieuws, maar het werd wel helder gepresenteerd. In de bijdrage van Zuzovsky die samen met Aitkin een paper had geschreven werd naast een multilevel benadering ook een lans gebroken voor interactie-effecten.

Uit de presentatie van Teddlie en Roberts bleek dat dezelfde set problemen binnen het effectiviteitsonderzoek, die in het vorige decennium uit analyses van Ralph en Fennessey onder andere naar voren kwam, nog steeds opgeld doet. Zo werden nu genoemd 'school change', 'teacher school relations', 'context issues', 'role of leadership', 'stability and consistency', 'curriculum issues', 'methodological issues', 'variance issues', 'theory existence/magnitude issues', 'district/state effects' and 'international studies'. Naar mijn mening is op deze groslijst van onderwerpen, waarvan er een aantal onderling meer overeenkomsten heeft dan verschillen, nogal wat kritiek uit te oefenen vanuit de overweging dat, wil men vooruitgang boeken, men ergens moet beginnen, bijvoorbeeld met de conceptualisering en fundamentele en funderende vraagstellingen, en dan op basis van datgene wat er aan onderzoek is verricht eens nagaan wat er zoal allemaal op het rooster is blijven liggen.

Op een overigens curieuze wijze werden in de groep van vooraanstaande onderzoekers op het terrein van 'school effectiveness and school improvement' drie groepen onderscheiden: de onderzoekers, de verbeteraars en degenen die een verbinding tussen deze twee terreinen zouden kunnen leggen.

Opvallend is dat er nauwelijks dwarsverbindingen zijn tussen onderzoek op het terrein van onderwijseffectiviteit en andere onderdelen van het AERA-programma die een basis voor verdere theorievorming en de praktijk van het onderzoek zouden kunnen leveren, zoals multilevelonderzoek, onderzoek naar 'opportunity to learn' als een mogelijkheid om beter grip te krijgen op datgene wat er in het curriculum gepresenteerd wordt (zie de curriculumissues in de lijst van Teddlie en Roberts) en ontwikkelingen op het terrein van instructie-effectiviteit, zoals onderzoek naar leerprocessen, nieuwe onderwijsvormen en instructievormen in de klas of verdere uitwerking van oude, zoals 'mastery learning'.

Vooral de bijeenkomst van de SIG mastery learning waarin het 25-jarig bestaan van mastery learning werd gevierd had iets 'tragisch'. In een zaal waar 500 mensen aanwezig konden zijn vierden 30 getrouwen het 25-jarig bestaan. En in de bijdrage van Guskey en Anderson werd duidelijk gemaakt dat 'learning for mastery' als instructiemodel duidelijk toekomst zou kunnen hebben. Omdat de resultaten ervan in diverse projecten bemoedigend zijn, werden verbindingen gelegd vanuit learning for mastery naar andere sectoren, zoals onderwijsverbeteringsprojecten en zelfs schooleffectiviteit. Edoch, er was (bijna) niemand om deze boodschap te beluisteren.

Ook de verdere uitwerkingen van de wijze waarop opportunity to learn gemeten zou kunnen worden, waarbij de kwaliteit van het onderwijsleerproces in beschouwing wordt genomen, zijn zaken die niet uitsluitend voor IEA-studies van belang zijn maar opgenomen zouden kunnen worden in onderzoek naar onderwijseffectiviteit en tot kwaliteitsverbetering van dit onderzoek zouden kunnen leiden.

Ten slotte 'school restructuring': dat is en blijft onduidelijk. Ook na het bijwonen van een aantal paperpresentaties op dit terrein blijkt dat allerlei 'onderzoekers' en auteurs daar verschillende zaken onder verstaan. Soms is het slechts een metafoor om een discussie in en over het onderwijs aan te zwengelen, soms is het een plan: meer doen met minder geld. In het beste geval vertoont restructuring naar mijn mening veel overeenkomst met school improvement, met een oriëntatie op effecten van onderwijs. Het speelt zich vooral op het niveau van de school en het onderwijsbeleid af, hoewel soms lijnen doorgetrokken worden naar de schoolorganisatie, (zoals het docententeam), en implicaties voor onderwijs op het niveau van de klas worden aangegrepen.

Het lezen van een vrij groot aantal publikaties dat op dit moment over restructuring verschijnt, zowel theoretisch als praktijkgeoriënteerd (Hoe mijn school te veranderen), maakt een en ander niet veel duidelijker. Misschien moeten we geduldig wachten tot ook deze waan weer voorbij is.

Divisie I: Education in the professions
(D. Dolmans, Rijksuniversiteit Limburg)

Deze divisie bevatte 24 thematische sessies over verschillende onderwerpen, waarvan een groot aantal betrekking had op medisch onderwijs en medische cognitie. Drie van de sessies die ik heb bijgewoond in deze divisie hadden betrekking op problem-based learning (PBL). In de Verenigde Staten is er op dit moment een grote toename van faculteiten die een PBL-curriculum introduceren. Ook in het Nederlandse onderwijs is een beweging waarneembaar van docent-gecentreerd onderwijs naar student-gecentreerd onderwijs. Deze toenemende aandacht voor student-gecentreerd onderwijs leidt er toe dat onderzoek naar effecten van innovatieve curricula op het leren van de student veel belangstelling trekt.

In de sessie genoemd 'A critical look at problem-based learning' werd uitgebreid verslag gedaan van een literatuurstudie betreffende de uitkomsten en implementatie van probleemgestuurd onderwijs aan medische faculteiten. Een aantal belangrijke conclusies hieruit werd door Albanese en Mitchell als volgt samengevat. In vergelijking met conventionele curricula wordt PBL hoog gewaardeerd door studenten. De prestaties van de afgestudeerden zijn vrijwel gelijk en soms zelfs beter, zoals blijkt uit klinische examens en evaluatiegegevens van faculteiten. Daarnaast blijkt dat studenten in PBL in vergelijking met studenten uit conventionele curricula meer gebruik maken van de bibliotheek voor het raadplegen van literatuur. Dit gedrag blijven ze ook vertonen tijdens de jaren waarin studenten aan kliniekstages bezig zijn. Echter, PBL-studenten scoren lager op de examens voor de basisvakken en zien zichzelf als minder voorbereid in de basisvakken. Albanese en Mitchell merkten op dat in verder onderzoek meer aandacht besteed moet worden aan de manier waarop de resultaten van innovatie curricula effectiever gemeten kunnen worden, aan de verschillen in implementatie van PBL, aan effectieve strategieën voor het verzekeren van content coverage en aan de lange termijn effecten. Verder onderzoek naar de vraag in hoeverre de gevonden effecten van PBL gegeneraliseerd kunnen worden naar vormen van PBL die kenmerken vertonen van conventionele curricula zou

eveneens aandacht verdienen.

In een tweetal andere sessies over PBL, geheten 'Problem-based learning: Evaluation of students and tutors' en 'Teaching and assessment issues in PBL' stond onderzoek naar de rol van de tutor centraal. De tutor vervult een belangrijke docentrol in probleemgestuurd onderwijs. Het belang van deze rol voor het leren van studenten vereist dan ook dat er aandacht besteed wordt aan de verdergaande professionalisering van tutoren. Een continue evaluatie van het functioneren van tutoren en hen voorzien van feedback daaromtrent vormt een van de onderdelen van de professionalisering van tutoren. Een tweetal hiervoor ontwikkelde instrumenten werd tijdens een van deze sessies gepresenteerd. Eén instrument werd ontwikkeld aan de McMaster University in Canada en een ander instrument werd ontwikkeld aan de Rijksuniversiteit Limburg. Alhoewel in de ene studie voornamelijk aandacht werd besteed aan de betrouwbaarheid van het instrument en in de andere studie vooral de validiteit en de bruikbaarheid centraal stonden, bleek dat beide instrumenten wat betreft opzet grote gelijkenissen vertoonden.

Daarnaast werd er tijdens deze sessies gediscussieerd over de noodzaak van 'expert' tutoren. Tutoren in PBL dienen vooral de ontwikkeling van de vaardigheden van studenten in actief en zelf-gestuurd leren, een van de belangrijkste doelen van PBL, te stimuleren. Studenten in probleemgestuurd onderwijs worden daarom aangespoord zelfstandig informatie te verwerven, een vaardigheid waarvan zij in hun latere beroep als medicus gebruik zullen moeten maken, zeker gezien de snel veranderende kennisontwikkeling in de geneeskunde. Dit impliceert dat studenten ertoe aangezet moeten worden zelf verantwoordelijkheid te nemen voor hetgeen ze leren, zodat ze kunnen aansluiten op hun eigen behoeften en interesses. Om dit te stimuleren dient de tutor vooral een begeleider van het groepsproces te zijn, waarvoor geen inhoudelijke deskundigheid is vereist. Om het leerproces te bevorderen stelt de tutor discussie stimulerende vragen. Hiervoor is wel een zekere mate van inhoudsdeskundigheid nodig. Uit onderzoek is gebleken dat expert tutoren een actievere rol vervullen in probleemgestuurd onderwijs dan niet-expert tutoren. Expert tutoren blijken langer aan het woord te zijn, geven meer direct antwoord op vragen van studenten en beïnvloeden in sterkere mate welke onderwerpen bediscussieerd worden. Dit gedrag van de tutor kan een negatieve invloed hebben op de ontwikkeling van vaardigheden in het zelf sturing geven aan het leerproces, aldus Wilkerson. Het probleem dat zich voordoet bij onderzoek naar verschillen in prestaties van studenten begeleid door expert tutoren en studenten begeleid door niet-expert tutoren, betreft de definiëring van expertise. In welke mate een tutor beschouwd kan worden als een 'expert' is namelijk niet alleen afhankelijk van de inhoudsdeskundigheid van de tutor, maar wordt ook bepaald door diens ervaring met de rol van tutor, het functioneren van onderwijsgroepen, groepsprocessen en de ontvangen trainingen.

In de sessie getiteld 'Studies in professional expertise' werd een vijftal studies gepresenteerd op het gebied van medische cognitie. De papers rapporteerden onderzoek over de wijze waarop expertise in complexe domeinen verworven wordt. Een van deze vijf studies betrof een onderzoek met als doel na te gaan of de ziektescripts van afgestudeerde geneeskunde studenten verschilden van die van experts. Uit de resultaten kwam naar voren dat experts meer profiteren van typische achtergrondinformatie en typische klachten en symptomen dan pas afgestudeerden. De mate van prototype van de casuïstiek blijkt dus van invloed te zijn op de kennisstructuur. Deze bevinding betekent dat geneeskunde studenten tijdens hun studie vooral ervaring dienen op te doen met typische patiënten in plaats van atypische patiënten. Een ander paper in deze sessie betrof onderzoek naar de ontwikkeling van expertise van economiestudent tot ervaren account. Uit de resultaten bleek dat naarmate het niveau van expertise toenam, er ook een toename waarneembaar was in diagnostische accuratesse, zoals ook werd verwacht. De ervaren accountants bleken bovendien in sterkere mate gebruik te maken van de gepresenteerde cijferinformatie, terwijl pas afgestudeerden ook verwacht werden deze cijferinformatie te kunnen interpreteren.

In de sessies in divisie I werd getracht de bevindingen op het terrein van de cognitieve psychologie te vertalen naar de onderwijspraktijk van alledag. Dit gold met name voor de bevinding

waarin geclaimd wordt dat onderwijs erop gericht moet zijn studenten aan te moedigen architecten te worden van hun eigen kennis en uiteindelijk zelf de volledige verantwoording voor hun leren te krijgen.

Divisie K: Teaching and teacher education
(N. Verloop, Rijksuniversiteit Leiden)

In 1989 omschreef ik in een zelfde Kroniek als deze Divisie K als "een betrekkelijke nieuwkomer in de lange lijst van AERA-divisies". Op dit moment kan vastgesteld worden dat deze divisie haar bestaansrecht meer dan bewezen heeft: zij is momenteel de op één na grootste divisie van de AERA. Dit hangt ongetwijfeld samen met het feit dat, in tegenstelling tot de organisaties die op Europees niveau op het gebied van de onderwijsresearch bestaan, het fenomeen 'onderwijzen' binnen de AERA vanouds een zeer centraal thema is geweest, zij het dat het oorspronkelijk binnen andere divisies aan de orde gesteld werd. Het is voor een Europees onderzoeker op dit terrein ook nu nog veel gemakkelijker om aansluiting te vinden binnen de Amerikaanse onderwijsresearch dan binnen de Europese traditie, die overwegend op leerprocessen bij leerlingen is gericht.

Ik zal trachten het enorme aantal op dit terrein gepresenteerde papers en symposia te ordenen in de vorm van een aantal trends. Uiteraard gaat het daarbij voor een deel om zaken die al enkele jaren in de belangstelling staan.

- Zoals het enkele jaren geleden gebruikelijk was om in de inleiding van een presentatie op dit terrein bij wijze van 'geloofsbelijdenis' zijn aanhankelijkheid te betuigen aan de teacher thinking-benadering, zo betuigden nu veel paperlezers vooraf hun steun aan het constructivisme. Dit, op zich uitermate zinnige, idee dat kennis mede door de gebruiker in de gebruikssituatie wordt gevormd, was alomtegenwoordig en werd er ook in contexten waar het nauwelijks functioneel was met de haren bijgesleept.
- In het kader van het onderzoek naar de lerarenopleiding was er veel aandacht voor de eigen ideeën en opvattingen (inclusief misconcepties) van de aanstaande leraar. Reeds een aantal jaren wordt op grote schaal het belang benadrukt van het aansluiten bij de subjectieve ideeën van de aanstaande leraar en het op de een of andere wijze in verband brengen van deze ideeën met de theoretische noties vanuit de opleiding. Concrete uitwerkingen van een dergelijke aanpak daarentegen zijn schaars. Wel is er een aantal pogingen gedaan deze impliciete theorieën, en veranderingen daarin, in kaart te brengen. Weber en Stronge rapporteerden een aardig onderzoek over veranderingen in de motivatie van aanstaande leraren onder invloed van de mentor, terwijl Mosenthal de mogelijkheden en gevaren van een portfolio-aanpak besprak voor het bereiken van een grotere bewustheid van de eigen opvattingen bij aanstaande leraren. Myers onderzocht, vanuit een verwante bedoeling, de bekendheid van aanstaande leraren met de 'objectieve' theorie.

  Omdat de onderzoekstraditie op dit terrein nog niet sterk is, trekt dit type onderzoek onvermijdelijk ook meer eenvoudige geesten aan, die dan veelal de volgende procedure volgen: men legt een vrij open vragenlijst voor aan een groep studenten of laat hen een logboek bijhouden. De aldus verkregen gegevens (omtrent 'impliciete theorieën', 'conceptions of teaching', 'images' of wat op dat moment ook maar de meest aansprekende aanduiding is) worden zonder enige structuur of theoretisch kader vrij associërend uitgesponnen, liefst geïllustreerd met ellenlange citaten uit studenten-logboeken. Wat het uitgangspunt of het doel van dergelijke excercities is blijft daarbij geheel duister.
- Met het (ook in Nederland niet onbekende) toenemende belang van de praktijkcomponent in de opleidingen is het niet verbazingwekkend dat er tal van bijdragen waren over de mogelijke varianten in de rol van mentor en van de praktijkschool als 'professional practice school'. Zo ging Ents bijvoorbeeld na in welke mate de aanstaande leraren, de mentoren en begeleiders vanuit de lerarenopleiding van de universiteit overeenstemming vertoonden over de taak van de laatstgenoemde. Een interessante bijdrage werd geleverd door Munby en Russel, die een nieuw concept, 'authority of experience', introduceerden. Zij vergeleken de verschillende varianten van 'autoriteit' die de aanstaande leraar in zijn schoolcarrière reeds is

tegengekomen en concludeerden dat deze de aanstaande leraar belemmeren om de eigen praktijkervaringen in de stageschool ten volle als bronnen van (authoritative) informatie te benutten. Munby en Russell benoemden daarmee naar mijn mening voor het eerst en op een originele wijze dit aspect van 'reflection-in-action' zoals dat door Schön is aangeduid. De hamvraag is vervolgens natuurlijk hoe de aanstaande leraren op de opleiding kunnen worden toegerust voor een dergelijke reflectie.

- Aan het thema 'reflectie' als zodanig waren wat minder sessies gewijd dan de voorgaande jaren, maar dat betekende niet dat dit geen centraal thema van veel bijdragen was. Het was eerder zo dat dit thema zodanig was ingeburgerd dat het belang ervan als vanzelfsprekend werd aangenomen. Zo spitste bijvoorbeeld het onderzoek naar verschillen tussen experts en beginners zich vaak toe op de vraag naar verschillen in reflectievaardigheid (zie volgende punt). Het was interessant om te zien dat er enige systematische aandacht is voor het gebruiken van cases als hulpmiddel bij het gericht bevorderen van reflectie in de lerarenopleiding (met bijdragen van Harrington en van Morine-Derschimer die, zoals altijd, met een gedegen effectonderzoek naar, in dit geval, het gebruik van cases kwam).
- Het expert-novice-onderzoek was ook dit jaar weer ruim vertegenwoordigd. Voor dit onderzoek geldt hetzelfde als voor het onderzoek naar subjectieve theorieën van leraren: velen werpen zich op dit onderwerp zonder voldoende theoretische en methodologische bagage, hetgeen resulteert in ellenlange structuurloze beschrijvingen van gevonden verschillen tussen experts en beginners, gelardeerd met al even uitgebreide citaten (Ponticell) of opsommingen van trivialiteiten of uiterst voor de hand liggende zaken (Blanton).

  Veel van het onderzoek naar expert-novice verschillen is nog beschrijvend van aard, hetgeen niet per se negatief gewaardeerd behoeft te worden, al vraagt men zich af hoe lang dit nog op deze wijze door kan gaan. Jensen en Templeton onderzochten de kenmerken waardoor enkele honderden 'Teachers of the Year' zich van de doorsnee leraar onderscheidden (en die vooral op het affectieve en motivationele vlak bleken te liggen), terwijl Cutler rapporteerde over datgene wat novice-leraren vooral in hun eerste jaar bezighield en dat, volgens haar, veel minder op louter het 'overleven' betrekking heeft dan tot nu toe werd aangenomen. Ze gebruikte daarvoor een interessante, unobtrusive wijze van meten, namelijk het analyseren van de inhoud van een computer-network dat door Harvard aan de aldaar afgestudeerde beginnende leraren ter beschikking was gesteld om onderling over hun ervaringen te kunnen communiceren. Uit een interessant onderzoek van Harris en Eggen bleek dat in het begin van het schoolpracticum de bekende verschillen tussen novices en experts op het terrein van het kunnen interpreteren van gebeurtenissen in de klas alleen nog maar toenamen, en wel vanwege het feit dat het de beginners ontbrak aan een interpretatiekader. Zij concluderen dat het loutere feit dat aanstaande leraren in de klas aanwezig zijn geen enkele garantie vormt dat zij zich meer in de richting van een expert zullen ontwikkelen. Aanstaande leraren hebben niet alleen 'modellen' nodig in de persoon van de mentor, maar eveneens een kader om dit gedrag te kunnen interpreteren en verantwoorde procedures om met 'hun' modellen over hun gedrag te communiceren. Ook Moore en Freeman toonden in een onderzoek aan dat, als er geen manieren gevonden worden om het denken van de expert (i.c. de mentor) voor de aanstaande leraar inzichtelijk te maken, het loutere feit dat men als aanstaande leraar in de klas verblijft nog niet leidt tot expert-gedrag (in dit geval: 'problem finding'). Het lijkt erop dat het doorwerken van dit soort inzichten in Nederland, waar een onberedeneerd vertrouwen in 'meer praktijk' de politiek lijkt te gaan bepalen, in de nabije toekomst cruciaal zal zijn voor de kwaliteit van de leraren die tot het beroep zullen worden toegelaten.
- Het, in de ogen van de West-Europese bezoeker, meest in het oog springende was de grote hoeveelheid aandacht die werd besteed aan het thema 'teacher assessment'. Tot voor kort had dit thema vooral betrekking op ervaringen met allerlei door de di-

verse staten verplicht gestelde, op accountability gerichte, evaluatie-systemen. Ook nu nog werd hierover gerapporteerd, variërend van gedegen studies naar de ervaringen van leraren daarmee en de, in dit geval als positief ingeschatte, gevolgen voor de professionele ontwikkeling (Lofton), tot enigszins onnozele, van geen enkel probleembewustzijn getuigende aankondigingen van de zegeningen van leraar-evaluatie (Silvestro). Het meest opvallend waren echter de door de federale overheid gestimuleerde initiatieven tot het ontwikkelen van meer omvattende en verantwoorde evaluatie-procedures. Ik memoreer hier een drietal verwante initiatieven waarover werd gerapporteerd:

(a) Educational Testing Service (te vergelijken met het Cito in Nederland), dat tot nu toe de National Teacher Examinations heeft verzorgd heeft het initiatief genomen tot de ontwikkeling van een geheel nieuw instrumentarium voor het beoordelen van beginnende leraren: 'The PRAXIS series', als ondertitel aangeduid met 'Professional Assessment for Beginning Teachers'. Door Dwyer en Villegas werd over de achtergronden van dit instrumentarium gerapporteerd. Opvallend daarbij is dat bij de evaluatie niet alleen observatie- en interviewgegevens zullen worden betrokken, maar ook de cognities van de beginnende docenten. Men tracht zich daarbij te baseren op hetgeen inmiddels bekend is uit teacher thinking onderzoek.

(b) De federale overheid heeft een enorme subsidie verstrekt aan het Center for Research on Educational Accountability and Teacher Evaluation (CREATE) te Kalamazoo. Hier zal, onder leiding van Stufflebeam, in een aantal meerjarige projecten getracht worden het evaluatie-instrumentarium voor docenten te verbeteren of te ontwikkelen. Zo is er een afzonderlijk project voor het ontwikkelen van beoordelingsmodellen (Scriven) en een project voor het ontwikkelen van procedures voor zelfevaluatie (Gullicksen).

(c) Naar aanleiding van de kritiek van o.m. de Cargenie Task Force ('A nation prepared') op het niveau van de Amerikaanse docent is de National Board for Professional Teaching Standards in het leven geroepen die zich bezighoudt met het ontwikkelen van criteria waaraan in functie zijnde docenten zouden moeten voldoen en met het (laten) ontwikkelen van evaluatie-instrumenten om dit te meten. Voor dat laatste is een budget van 50 miljoen dollar beschikbaar. Bij dit ontwikkelingswerk zijn tal van bekende onderzoekers betrokken, waaronder Shulman. Een speciale 'Technical Analysis Group' (onder leiding van Jaeger) houdt zich bezig met de methodologische aspecten, m.n. de generaliseerbaarheid.

Het lijkt, gegeven de recente ontwikkelingen in Nederland op dit terrein, zoals de in september 1993 verschenen kabinetsreactie op het rapport van de Commissie Van Es ('Vitaal Leraarschap'), van belang deze ontwikkelingen nauwgezet te volgen.

Aan deze Kroniek werkten mee: *J.J.H. van den Akker, B.P.M. Creemers, S. Dijkstra, D. Dolmans, T. de Jong, J.J.G. van Merriënboer, P. Sanders, N. Verloop*

De eindredactie werd verzorgd door *N. Verloop*

Boekbespreking

A. J. J. M. Ruijssenaars
Rekenproblemen. Theorie, diagnostiek, behandeling
Lemniscaat, Rotterdam 1992, 294 pagina's ƒ 55,- ISBN 0 6069 857 6

Het is al weer ruim een kwart eeuw geleden dat met de diagnose 'discalculie' de problemen van de zwakke rekenaar in kaart werden gebracht. Inmiddels zijn er over het leren, denken, handelen en probleemoplossen van kinderen nieuwe theorieën ontwikkeld en Ruijssenaars wil in zijn boek nagaan wat de betekenis is van deze, maar ook van enkele 'gevestigde', theorieën voor de diagnostiek en behandeling van rekenproblemen.

De auteur heeft voor zijn publikatie dan ook een brede opzet gekozen die begint met een uiteenzetting van enkele opvattingen over leren en leerproblemen. In het voetspoor van Dumont worden primaire en secundaire leerproblemen onderscheiden. Primair betekent dat er iets mis is met de cognitieve ontwikkeling van het kind, terwijl een kenmerk van secundaire problemen is dat er (onder meer) sprake is van een intelligentietekort. Aangezien intelligentie een kwestie is van cognitieve ontwikkeling is het niet geheel duidelijk waarop Dumont zijn indeling in primaire en secundaire leerproblemen baseert.

Vervolgens wordt nagegaan wat in de literatuur zoal onder rekenen wordt verstaan en de conclusie luidt dat rekenen en wiskunde, "niet volledig zijn te scheiden" (p. 39). Dat is juist en in de vakdidactiek wordt die scheiding zelfs al geruime tijd bestreden. Rekenen werd ten onrechte een vak apart als gevolg van een decennia lang op algoritmisering en cijferen gericht koopmansrekenen. Ook Ruijssenaars onderkent het onvruchtbare van die scheiding, want zo zegt hij "Het wiskundig inzicht ligt ten grondslag aan het rekenkundig inzicht" (p. 39). Toch klinkt in het boek die laatste opvatting niet steeds even luid en duidelijk door en zullen wiskundigen trouwens direct opmerken dat het kunnen opereren met allerlei soorten getallen (ordinaalgetal, numeriek getal, meetgetal, gebroken getal, negatieve getallen etc.) wiskundig inzicht veronderstelt. Terecht benadrukt de auteur dat het (leren) rekenen een actief - tegenwoordig zegt men ook wel een constructief - proces is.

De vraag wat rekenproblemen zijn, wordt vanuit verschillende standpunten bekeken. Vanuit de handelingstheorie, vanuit de informatieverwerkingstheorie en de zogenoemde probleemoplossingsbenadering. Waarom juist deze theorieën zijn gekozen en niet andere (de vakdidactische van de vroegere Wiskobasgroep ontbreekt bijvoorbeeld) maakt de auteur niet duidelijk.
Wat is getalbegrip? "Getalbegrip is te beschouwen als *de* basis waarop *het* rekenen berust" (p. 64, curs. J.N.). Terecht wordt in dit verband het belang van tellen onderstreept, maar *het* rekenen berust op meer dan getalbegrip, denk maar aan verhoudingsbegrip, maatbegrip, inzicht in het gebruik van handige strategieën. In navolging van Dumont - en deze baseert zich weer op Piaget - wordt conservatie als een basisvoorwaarde voor getalbegrip beschouwd. Gelukkig bespreekt de auteur ook onderzoekers die dit standpunt bekritiseerden (helaas niet het onderzoek van Donaldson). De vraag blijft toch steeds wat nu precies het belang van het conserveren is voor het leren rekenen. Piaget onderzocht het conservatiefenomeen in eerste instantie - *niet* als rekenvoorwaarde maar - om de door volwassenen onbeïnvloede spontane ontwikkeling van het kinderlijk denken te kunnen bestuderen. Ruijssenaars is terecht van mening dat de 'voorwaarden' zich tijdens het proces van leren rekenen kunnen ontwikkelen, waarmee het dus in eigenlijk zin geen voorwaarden meer zijn.

"Rekenen als proces" is de titel van het hoofdstuk waarin het proces van het leren rekenen in algemene zin vanuit de bovengenoemde drie theorieën wordt besproken. De leertheorie van Gal'perin krijgt uitvoerig aandacht en in dat verband wordt op het belang gewezen van het controleren en evalueren (kortom het reflecteren) van kinderen. Behartenswaardig is de opmerking van de auteur dat de leeromgeving sterk moet zijn "op de punten waarin de leerling zwak is" (p. 103). De A.T.I.-onderzoekers hebben jarenlang het tegenovergestelde verdedigd, namelijk dat je het onderwijs moet afstemmen juist op de sterke karakteristieken van de leerlingen.

Een in de orthopedagogische vakliteratuur vaak terugkerend discussiepunt luidt welke samenhang er bestaat tussen leren rekenen en psychologische variabelen, in het bijzonder die varaibelen die men op basis van standaardintelli-

gentietests op het spoor tracht te komen. De auteur stelt vast dat prestaties op rekentests en op intelligentietaken "tot op zekere hoogte iets gemeenschappelijks hebben" (p. 144), een conclusie waarmee hij weinig tegenspraak zal uitlokken.

Een belangrijk hoofdstuk is het zesde waarin de diagnostiek van rekenproblemen wordt behandeld. De auteur presenteert een 'totaaloverzicht', bedoeld als ordening van 'domeinen' waarin zich problemen kunnen voordoen. Daarna volgt een bespreking van verschillende onderzoekstypen variërend van formeel-incidenteel tot informeel-continu. Ruijssenaars merkt hier op dat Gal'perin de theorie van Piaget zou hebben uitgewerkt. Gal'perin heeft echter Piaget, voor zover ik weet, alleen maar bestreden (met name Piagets constructivisme en het idee van de fasengewijze ontwikkeling).

De auteur bespreekt kritisch het diagnostisch materiaal van Van Erp: 'Er wordt niet gevraagd hoe het kind werkt, maar "of" het de enig juiste manier kent'. De kritiek van Vuurmans echter op hetzelfde materiaal gaat de auteur weer te ver. Vuurmans denkt namelijk vanuit een wiskundig en 'eveneens volledigheid claimend' standpunt. Ik zei het al: het onderscheid dat wordt gemaakt tussen rekenen en wiskunde is niet houdbaar. Het mathematiseren begint zelfs al op jonge leeftijd, zij het op elementair niveau, namelijk bij het tellen (akoestisch, synchroon, structurerend en verkort tellen).

In het kader van het informele continue onderzoek geeft Ruijssenaars enkele belangrijke suggesties voor observatie zoals: het laten reconstrueren van procedurele kennis en nagaan of een kind gemakkelijk van werkwijze kan wisselen.

Een aparte beschouwing wordt gewijd aan het psychologisch onderzoek. De auteur waarschuwt terecht dat de interventie (de behandeling dus) van intelligentietests niet erg profiteert. Waarom dan toch die aandacht voor intelligentieonderzoeken? In dit verband wordt gewezen op het "logische karakter van het rekenen"; rekenen is een kwestie van "omgaan met logische structuren" (p. 196), die men ook in intelligentietests tegenkomt. Natuurlijk, er vinden in het reken/wiskundeonderwijs voortdurend formaliseringsprocessen plaats die leiden tot het gebruik van mathematische middelen op steeds hoger niveau. De logica speelt daarbij onmiskenbaar een rol. Het mathematiseringsproces begint echter niet met het hanteren van logische structuren, maar met het hanteren van *eigen* middelen, bij de *eigen* 'logica van het kind', om het zo te zeggen. Dat is een andere logica dan die van de intelligentietests. Ruijssenaars lijkt hier te blijven steken in een een-dimensionale beschrijving van veelal algoritmisch verlopende reken-denkprocessen. Rekenen/wiskunde is niet alleen een kwestie van abstractie maar ook van concretiseren, van voorstelbaar maken, van experimenteren, van blikwisseling, context en modelgebruik.

De visie op rekenen/wiskunde waar een onderzoeker van uitgaat, klinkt altijd sterk door in de benadering van rekenproblemen van kinderen. Boris een kind uit de 4e groep (p. 198) blijkt niet in staat het geleerde toe te passen. Dit wordt veroorzaakt - zo vermoedt de auteur - door onvoldoende logisch inzicht. Het is echter uit onderzoek bekend dat als kinderen niet van meet af aan leren toepassen, ze dit nadien ook niet goed kunnen. De geleerde procedures worden geïsoleerd aangeleerd en functioneren dus niet. En juist het toepassen (in contexten) is wat in de traditionele didactiek ontbreekt, gesteld dat Boris volgens die didactiek leerde rekenen. Boris bleek overigens over een normale intelligentie te beschikken en dus - zo luidde het advies - kon met een gerust hart een beroep worden gedaan op de verbale mogelijkheden en het denken van de jongen. Natuurlijk kan dit geen kwaad, maar of hij daardoor beter zal leren toepassen, is de vraag.

In het laatste hoofdstuk wordt de behandeling van rekenproblemen besproken, uitgaande van enkele belangrijke (maar algemene) instructieprincipes. Interessant zijn de verschillen die Ruijssenaars bespreekt tussen de vakdidactische principes en de orthodidactische benadering. Zo wordt de laatste omschreven als 'sturend' om het kind zodoende een veilige basis te bieden. Op dat punt kan de orthodidactiek de realistische vakdidactiek nog wat leren en het is daarom jammer dat dit idee niet uitvoeriger wordt uitgewerkt.

De behandeling van rekenproblemen moet uitgaan van een theorie en het zijn weer (onder meer) de bovengenoemde drie theorieën die in dit verband worden geanalyseerd. Die theorieën worden gewogen en in het algemeen zwaar genoeg bevonden.

Het boek van Ruijssenaars voert als titel "Re-

kenproblemen". Die titel is nogal ruim want in feite worden voornamelijk het rekenen onder de 100 en de cijferalgoritmen besproken. Een aantal gebieden blijven vrijwel geheel buiten beschouwing (zoals meten, verhoudingen, procenten, grafieken, schatten). Wat heeft het boek te bieden voor (ortho)didactici die kinderen met rekenproblemen onderzoeken? In elk geval een diversiteit aan beschouwingen en analyses. Ruijssenaars stelt zich daarin echter nogal afstandelijk op. Het is, om een voorbeeld te geven, uit onderzoek bekend welke problemen kinderen ondervinden bij het optellen en aftrekken onder de 20. Deze kinderen hebben namelijk sterk de neiging *alle* opgaven *tellend* op te lossen, omdat ze zich bij deze aanpak veilig voelen. In zulke gevallen moet onderzocht worden of een kind in staat is de telmethode te vervangen door effectievere procedures zoals: verkort en structurerend tellen en het gebruik maken van getalbeelden. Met dit type onderzoek blijft men 'dicht bij huis': dicht bij het kind en dicht bij het rekenen. Ik zie hier meer in dan in het zoeken naar stoornissen in de intelligentie, in de informatieverwerking of in het korte en lange termijngeheugen.

Een uiteindelijk oordeel over het hier besproken boek is niet zo gemakkelijk te geven. Dat hangt er maar van af voor wie het boek is bedoeld. Het biedt veel informatie, er worden veel theorieën en benaderingen besproken, er zijn nuttige tips uit te halen en het biedt ook genuanceerde analyses van bestaande materialen. Aan zulk een boek kan in de opleiding behoefte bestaan, al zal de docent voor kritisch begeleiding moeten zorgen.

Voor de praktijk biedt het wel enige, maar niet veel nieuwe, handvatten. Een probleem in dit verband is, lijkt me, dat op ruim driekwart van de basisscholen inmiddels didactisch vernieuwde - namelijk op realistische leest geschoeide - programma's worden gebruikt, terwijl Ruijssenaars grotendeels nog uitgaat van een deels traditioneel, deels structuralistische visie op reken/wiskundeonderwijs.

Onderzoekers en andere deskundigen op het gebied van de vakdidactiek rekenen/wiskunde zullen vermoedelijk over vele kwesties die besproken worden met de auteur in debat willen treden. Een eerste aanzet tot dit debat is bij deze wellicht gegeven.

J. M. C. Nelissen

Mededelingen

ORD 1994

De Onderwijsresearchdagen 1994, georganiseerd door de vakgroep Onderwijskunde van de Universiteit Utrecht, onder auspiciën van de VOR, zullen worden gehouden van 24 t/m 26 mei 1994 in de Jaarbeurs te Utrecht. De informatiefolder, evenals inschrijvingsformulieren en aanmeldingsformulieren voor presentatievoorstellen, zijn op te vragen bij het FBU Congresbureau/Universiteit Utrecht, t.a.v. mevr. L. Donkers, Heidelberglaan 8, 3584 CS Utrecht, tel.: 030-535044/1642; fax: 030-533667

Nederlands Psychologencongres 1994

Op 20 en 21 oktober 1994 wordt voor de achtste keer het Nederlands Psychologencongres georganiseerd. Het congres, met als thema 'Psychologie bij Bewustzijn' vindt plaats op de Katholieke Universiteit te Nijmegen, onder auspiciën van het NIP en de Stichting voor Gedragswetenschappen (voorheen PSYCHON). Doel van het congres is de bevordering van de wisselwerking tussen psychologische praktijk en wetenschapsbeoefening.

De congrescommissie roept belangstellenden op tot aanmelding van bijdragen. Gevraagd worden voorstellen voor 1. symposia (bestaande uit 3 à 4 lezingen), 2. individuele papers (worden gecombineerd tot papersessies of postersessies), 3. trainingen, 4. audiovisuele of computerdemonstraties. Voorstellen voor symposia en individuele papers dienen vóór 1 maart 1994 aan het congressecretariaat te worden gezonden, waartoe u zich ook kunt wenden voor verdere informatie. Het adres luidt: Congressecretariaat. NIP-bureau, mevr. drs. M.A. Breijer, Postbus 9921, 1006 AP Amsterdam, tel.: 020-6109596; fax: 020-6108361.

Inhoud andere tijdschriften

Pedagogisch Tijdschrift
18e jaargang, nr. 4, 1993

Pedagogisch commentaar, door K. Potemans
Loopbaanverhalen en professionele ontwikkeling. Een narratief-biografisch onderzoek bij leerkrachten basisonderwijs, door G. Kelchtermans
Besprekingsartikel: Charles Taylor en de hernieuwde aandacht voor het authentieke individu, door P. Smeyers
Kroniek: Een crosscultureel historisch-pedagogisch congres ISCHE XV, door G. Boone

Tijdschrift voor Onderwijsresearch
18e jaargang, nr. 3, 1993

Themanummer De ontwikkeling van expertise in verschillende domeinen

De ontwikkeling van expertise in verschillende domeinen: inleiding, door H.G. Schmidt
Kennissencapsulatie, door H.G. Schmidt en H. Boshuizen
Probleemoplossen, leren en onderwijzen in exacte vakken: een voorbeeld uit de natuurkunde, door T. de Jong en M.G.M. Ferguson-Hessler
Kennisstructuur en hypothesevorming; verschillen tussen beginners en experts in de geneeskunde, door H.P.A. Boshuizen, P.P.M. Hobus, E.J.F.M. Custers en H.G. Schmidt
Meta-knowledge and the development of expertise in complex domains, door M.G.M. Ferguson-Hessler

Tijdschrift voor Orthopedagogiek
32e jaargang, nr. 10, 1993

Schoolbegeleiding in de toekomst, door R. de Groot
Toekomstige verschijningsvormen van onderwijsbegeleiding, door W. Meijer
Het meten van de opvoedingsomgeving met de Vragenlijst Meegemaakte Gebeurtenissen, door J.W. Veerman en L.T. ten Brink
Van impliciete kennis naar expliciete criteria: De Heerlense Ernst Taxatie Schaal (HETS), door W.J.P.J.M. Steerneman, H.J. Pelzer en E.E.J. de Bruyn

Ontvangen boeken

Adriani, P., *Overbruggen van verschillen. Hulpverlening aan jeugdigen uit allochtone groepen.* Sectie Interculturele Pedagogiek, Rijksuniversiteit, Leiden, 1993, *f* 24,50.

Angenent, H., *Opvoeding en persoonlijkheidsontwikkeling* (6e, herziene druk). Intro, Nijkerk, 1993, *f* 44,50.

Boer, B. de, Reubel, F., Reinards, R., & Sanden, J. van der, *Zelfstandig leren in beroepsopleidingen. Meer kansen op de Europese arbeidsmarkt.* Wolters-Noordhoff, Groningen, 1993, *f* 39,50.

Boonstra, O.W.A., *De waardij van eene vroege opleiding. Een onderzoek naar de implicaties van het alfabetisme op het leven van inwoners van Eindhoven en omliggende gemeenten, 1800-1920* (dissertatie). Landbouw Universiteit, Wageningen, 1993.

Breeuwsma, G., *Alles over ontwikkeling. Over de grondslagen van de ontwikkelingspsychologie* (dissertatie). Boom, Meppel, 1993, *f* 48,50.

Broeder, P., Extra, G., Habraken, M., Hout, R. van, & Keurentjes, H., *Taalgebruik als indicator van etniciteit.* Tilburg University Press, Tilburg, 1993, *f* 29,–.

Brouwers, H., & Goor, H. van, *Diagnostiek en behandeling van leesproblemen* (3e herziene druk). Intro, Nijkerk, 1993, *f* 39,50.

Ehrlich, K., Rehbein, J., Thije, J.D. ten (red.), *Kennis, taal en handelen.* Van Gorcum, Assen, 1993, *f* 49,50.

Goffree, F., *Kleuterwiskunde.* Wolters-Noordhoff, Groningen, 1993, *f* 39,50.

Graaf, F. de, Keulen, A. van, Projectbureau Mutant (red.), *Interculturee! opvoeden.* Uitgeverij SWP, Utrecht, 1993, *f* 32,50.

Hartensfeld, Fr., *De mystiek van het denken. De betekenis van het dialogisch principe in het denken van Martin Buber.* Gooi & Sticht, Baarn, 1993, *f* 22,50.

Hoop, F. de, & Janson, D.J., *Omgaan met (hoog)begaafde kinderen. Een andere kijk op (hoog)begaafdheid in school en gezin.* Intro, Nijkerk, 1993, *f* 29,75.

Jong, F.P.C.M. de, *Zelfstandig leren. Regulatie van het leerproces en leren reguleren; een procesbenadering* (dissertatie). Katholieke Universiteit Brabant, Tilburg, 1992.

Leij, A. van der & Kappers, E.J. (red.), *Zorgverbreding. Bijdragen uit speciaal onderwijs aan basisonderwijs* (5e, herziene druk). Intro, Nijkerk, 1993, *f* 49,50.

Lens, W., *Studiemotivatie. Theorie voor de praktijk op school en thuis.* Universitaire Pers, Leuven, 1993, BFrs. 495,–.

Matthijs, M., *Mythe van de jeugd.* 1. *Cultuur en leefwereld,* 2. *Beleid en interventie.* Wolters-Noordhoff, Groningen, 1993, *f* 42,50, resp. *f* 39,50.

Mönks, F. & Ypenburg, I., *Hoogbegaafde kinderen thuis en op school.* Dekker & Van de Vegt, Assen, 1993, *f* 17,50.

Mönnink, J. de (eindred.), & Ogg, F., *Opvoeden tot verantwoordelijkheid. Waardenoriëntatie en de pedagogische opdracht in onderwijs.* Kok Educatief, Kampen/K.P.C., 's-Hertogenbosch, 1993, *f* 49,50.

Muller, L., *Directieve thuisbehandeling.* Uitgeverij SWP, Utrecht, 1993, *f* 34,50.

Pouwelse, W.J., *Haar verstand dienstbaar aan het hart. Middelbaar onderwijs voor meisjes, debatten, acties en beleid 1860-1917* (dissertatie). Uitgeverij Het Spinhuis, Amsterdam, 1993, *f* 37,50.

Reezigt, G.J., *Effecten van differentiatie op de basisschool* (dissertatie). RION, Groningen, 1993.

Schonewille, B., *Heterogeniteit, leerkrachttijd en de rol van individuele aandacht in het speciaal onderwijs* (dissertatie). Vrije Universiteit, Amsterdam, 1993.

Smuling, E.B., Hout, J.F.M.J. van, Mirande, M.J.A., *Colleges en presentaties.* Wolters-Noordhoff, Groningen, 1993, *f* 29,50.

Tomic, W., & Span, P. (red.), *Onderwijspsychologie. Beïnvloeding, verloop en resultaten van leerprocessen.* Uitgeverij Lemma, Utrecht/Open Universiteit, Heerlen, 1993.

Vreeburg, B.A.N.M., *Identiteit en het verschil. Levensbeschouwelijke vorming en het Nederlands voortgezet onderwijs* (dissertatie). Boekencentrum Zoetermeer BV., Zoetermeer, 1993, *f* 59,–.

Vries, G. de, *Het pedagogisch regiem. Groei en grenzen van de geschoolde samenleving* (dissertatie). J.M. Meulenhoff BV., Amsterdam, 1993, *f* 34,50.

In 1993 werd redactionele medewerking verleend door:

C. A. J. Aarnoutse (K.U. Nijmegen)
R. van den Berg (K.U. Nijmegen)
M. Boekaerts (R.U. Leiden)
K. Bügel (Cito Arnhem)
H. Dekkers (I.T.S. Nijmegen)
J. Donkers (U.v.A. Amsterdam)
J. Elen (K.U. Leuven)
M. Elshout-Mohr (S.C.O. Amsterdam)
K. de Glopper (S.C.O. Amsterdam)
G. Haselager (K.U. Nijmegen)
A. van der Hoeven - van Doornum (I.T.S. Nijmegen)
R. Hoksbergen (U.U. Utrecht)
D. Janssen (U.U. Utrecht)
P. Jungbluth (I.T.S. Nijmegen)
J. Kloprogge (Min. O&W Zoetermeer)
F. A. J. Korthagen (U.U. Utrecht)
N. A. J. Lagerweij (U.U. Utrecht)
J. M. G. Leune (E.U.R. Rotterdam)
G. J. Meesterberends-Harms (RION Groningen)
M. J. C. Mommers, Tilburg
J. Scheerens (U.T. Enschede)
T. Vallen (K.U.B. Tilburg)
S. A. M. Veenman (K.U. Nijmegen)
J. Vermunt (K.U.B. Tilburg)
W. L. Wardekker (V.U. Amsterdam)
A. M. L. van Wieringen (U.v.A. Amsterdam)
J. M. A. M. Willems (K.U. Nijmegen)
J. M. Wijnstra (Cito Arnhem)